Psychometrika

CONTENTS

THE MEASUREMENT OF SPEED IN MOTOR PERFORMANCE - - - - - - - - - - - - - - 173
C. H. McCLOY

WEBER LAWS, THE WEBER LAW, AND PSYCHOPHYSICAL ANALYSIS - - - - - - - - - - - 183
ALSTON S. HOUSEHOLDER AND GALE YOUNG

SCORING WEIGHTED MULTIPLE KEYED TESTS ON THE IBM COUNTING SORTER - - - - - - - - 195
RALPH BEDELL

CONTRIBUTIONS TO THE MATHEMATICAL THEORY OF HUMAN RELATIONS III - - - - - - - - 203
N. RASHEVSKY

THE EVALUATION OF MULTIPLE AND PARTIAL CORRELATION COEFFICIENTS FROM THE FACTORIAL MATRIX - - - - - - - - - - - - - 211
P. S. DWYER

VOLUME FIVE SEPTEMBER 1940 NUMBER THREE

PSYCHOMETRIKA—VOL. 5, NO. 3
SEPTEMBER, 1940

THE MEASUREMENT OF SPEED IN MOTOR PERFORMANCE*

C. H. McCLOY
STATE UNIVERSITY OF IOWA

When the centroid method of factor analysis was applied to two sets of data on athletic performances, three significant factors emerged: strength, velocity, and dead weight. Scores on this speed factor were predicted by the multiple regression technique, the factor loadings on the speed factor being used as the criterion correlations, and these predicted scores were correlated with each of the other variables. When the original tables, augmented by the new speed variable, were refactored, the computed speed factor fell on the speed axis as a primary trait. It is thus shown that it is possible to isolate and measure a factor which appears in variables under consideration only as a compound.

In the analysis of many of the simpler motor skills used in physical education and athletics, such as in the runs, the jumps, and the throws, three orthogonal factors usually emerge, which are readily identifiable as (1) *strength,* or the force of contraction of muscular tissue, (2) *speed* of movement, and (3) *dead weight.* For the purposes of preliminary orientation, it may be well to discuss these factors briefly before considering the results of the study presented in this paper.

In research in this type of motor activities, it is quite practicable to insert specific variables of relatively known composition. For example, there are available numerous tests of strength which involve no significant speed component but which involves simply slow, forceful contractions. They are so simple that a few rehearsals suffice to enable the subject to master the skill; hence, educability in this specific performance is not an appreciable problem. If a number of such tests are used in a factor study, they invariably cluster together in factorial space. (For example, see Table II, variables 6 to 9.)

Speed, on the other hand, is not so readily definable. It is speed of movement and is orthogonal to strength. It must not be confused with *power.* In such events as a high jump, the velocity of projection upward from the ground is a resultant of the *power* expended. This power (P) is the product of force (F) and velocity (V), ($P = FV$). We are, therefore, used to associating our concept of speed with this

* Much of the expense of the statistical computations connected with this study was financed by a grant from the Carnegie Foundation for the Advancement of Teaching.

resultant of the expenditure of power. This may be made clearer with an illustration. Suppose we were to choose 1000 men of the same age, height, weight, and strength. They would not all respond, in motor athletic performance, with the same speed. Residuals from predicted velocity in actual samples (predicted in adults from height, weight, and strength) distribute themselves approximately normally. It is these differences in speed that are measured by this speed factor; and they may be due to differences of viscosity in the sarcoplasm of the muscular tissue, or they may be due to differences in leverage in the attachments of the muscles.

For example, in one individual, the distance from the olecranon process of the elbow to the center of the joint with the humerus is 1.7 cms. The distance from this point to the distal end of the ulna is 26.3 cms. This relationship of the power arm to the work arm differs significantly in different men, resulting in greater speed at the distal end of the forearm with a given speed of movement at the end of the power arm. Whatever the cause, the result is measurable through a factor analysis.

The third factor is *dead weight*. From an athletic point of view, weight is divisible into two parts. "Live weight" is that weight which is muscular and which pays its way. It is indirectly measurable through the *strength factor*. This, in the adult male, makes up about 43 per cent of the total. "Dead weight" is the remainder—the freight that is carried by the muscles. This dead weight is *positively* correlated with weight-throwing events such as the shot put, for it aids through its inertia in resisting the reaction of the throw. It is frequently *negatively* correlated with the runs and jumps, for here it is just extra baggage.

Since these three factors are so universally found in motor events, variables calling for them can easily be included in any appropriate battery of motor tests in order to isolate and identify these factors from among all of the factors that may emerge in the more complicated motor events.

Since the track and field or power type of events falls on a plane that is oblique to that of the strength factor, it is plain that these events are compounds of strength and speed. Hence, we wished to isolate this speed factor for experimental purposes.

In the present study, two sets of data were used. The first consisted of athletic performances of 100 college men for the following events:

4-pound shot
8-pound shot

12-pound shot
16-pound shot
24-pound shot
Putting a baseball (in the same manner as for the shot)
The unweighted strength index (total strength)
The "athletic" or weighted strength index (has some dead weight)
The Sargent Jump (springing straight upward)
Weight

These data were collected by Mr. James W. Coleman* of the University of Nevada. The ages of the group ranged from 18 to 24.

The second set of data consisted of athletic performances of 163 junior-high-school girls in the following events:

A short dash
The standing broad jump
Basketball throw for distance
3-pound shot put
12-pound shot put
Back strength
Leg strength
Shoulder strength
Grip strength
Total strength
The Sargent Jump
Weight

These data were collected by Miss Jane E. Harris† of Sioux City, Iowa. In both studies, the shot-putting events were done from a stand, without the usual preliminary hop.

These data were factored and rotated by the "two-at-a-time" method, and three significant factors emerged, which were readily identifiable as the three mentioned above—strength, velocity, and dead weight. The rotated factor loadings are given in Tables I and II in roman type at the upper left.

We then used multiple regression equations to predict the speed factor, using the factor loadings on the speed factor for the correlations with the criterion, and the intercorrelations between the variables for the other correlations (see Tables III and IV).

The resultant multiple regression equations for the speed factors were as follows (the results are in terms of *T*-scores):

* Coleman, James W. The differential measurement of the speed factor in large muscle activities. *Research Quarterly*, 1937, 3, 123-130.

† Harris, Jane E. The differential measurement of force and velocity for junior high school girls. *Research Quarterly*, 1937, 4, 114-121.

For the men:

Speed factor = 1.573 (4-pound shot in ft.) + .308 (Sargent Jump in cms.) — .0185 (Strength Index in lbs.) — 2.512;
$R = .887$.

For the girls:

Speed factor = .824 (Standing broad jump in in.) + .893 (3-pound shot in ft.) — .0505 (Strength Index in lbs.) — 7.00;
$R = .909$.

It will be noted that in the formulas above, the Strength Index,* which has approximately a zero correlation with the speed factor and with the dead weight factor, was used to partial out the strength in the power events used.

Since a factor has no absolute mean or standard deviation,† we used 50 for the mean and 10 for the standard deviation, and predicted the factor in the form of a *T*-score.‡ The predicted speed score was found for each subject and correlated with each of the other variables, and the material was refactored, yielding the centroid factor matrices in Tables V and VI.

The resulting rotated factor loadings are found in Tables I and II in italics at the lower right. In determining the number of factors to which to carry the results, two criteria were used. The fourth factor was computed and plotted against each of the others. The clusters seemed very slender along the axis and were entirely meaningless to the experimenter. Second, the third factor residuals were surveyed, and in view of the number of cases, seemed not to justify further computations. The distributions of the absolute values of the residuals were as follows:

* The "Strength Index" is a total strength score made up of the sum of a number of strength test variables. See McCloy, C. H. *Tests and Measurements in Health and Physical Education.* New York: F. S. Crofts, 1939, Chapter 4.

† All of the other variables used have absolute values, regardless of the group studied. Thus a 10-year old boy and a 20-year old man, each of whom runs 100 yards in 12 seconds, achieve the same record. The scores for a speed *factor*, however, computed for each of the two age groups, would probably be different since they would be in terms of standard deviations from means of different experimental groups.

‡ We prefer *T*-scores to *z*-scores (or standard scores) because we find them to be much more meaningful to students, who frequently seem to have a high natural resistance against minus signs.

	Men	*Girls*
.16	2	
.14	1	
.12		
.10		1
.08	1	4
.06	6	6
.04	7	17
.02	10	27
.00	28	23

These are the third factor residuals of the second process of factoring. The results are interesting. It will be noted that (Tables I and II):

TABLE I

Rotated Factorial Matrix
College Men

	I Strength factor	II Speed factor	III Dead weight factor	*h*
	.480	.770	.077	.910
1. 4-lb. shot put	*.395*	*.850*	*.177*	*.954*
	.489	.641	.476	.936
2. 8-lb. shot put	*.532*	*.589*	*.455*	*.915*
	.500	.611	.458	.913
3. 12-lb. shot put	*.477*	*.501*	*.611*	*.923*
	.461	.703	.468	.962
4. 16-lb. shot put	*.445*	*.583*	*.633*	*.969*
	.517	.521	.491	.883
5. 24-lb. shot put	*.494*	*.420*	*.651*	*.919*
	.323	.556	.186	.669
6. Baseball put	*.262*	*.551*	*.250*	*.660*
	.936	.000	.000	.936
7. Strength index	*.949*	*.030*	*.004*	*.950*
	.890	—.021	.241	.922
8. Athletic strength index	*.978*	*.228*	*—.136*	*1.013*
	.288	.577	—.326	.722
9. Sargent Jump	*.064*	*.684*	*—.063*	*.690*
	.653	.010	.431	.782
10. Weight	*.703*	*—.022*	*.317*	*.771*
	.....	.....	.....	.....
11. Computed Speed factor	*.082*	*.930*	*—.003*	*.934*

Figures in roman type, upper left of each space, are for the first factorial matrix. Figures in italics, lower right of each space, are for the second factorial matrix. The computed speed factor was computed from the first factorial matrix and hence was not included in it.

(1) The computed speed factor falls on the speed axis and has a correlation of .0 with the strength and the dead weight factors.

(2) Its vector is approximately the same as the multiple correlation with the factor, as would be expected theoretically. The minor departures from the theoretical values are readily accounted for by the differences in the assumed values in the diagonals in correlation matrices of this size. The Strength Index measures the strength factor in almost pure form.

TABLE II

Rotated Factorial Matrix
Junior High School Girls

	I Strength factor	II Speed factor	III Dead weight factor	*h*
	.279	.642	—.150	.716
1. Dash	*.347*	*.573*	*—.101*	*.677*
	.327	.775	—.265	.882
2. Standing broad jump	*.321*	*.854*	*—.256*	*.947*
	.557	.613	.190	.850
3. Basketball throw for distance	*.610*	*.543*	*.250*	*.853*
	.555	.602	.139	.830
4. 3-lb. shot put	*.578*	*.607*	*.171*	*.855*
	.671	.421	.352	.867
5. 12-lb. shot put	*.700*	*.400*	*.347*	*.877*
	.823	.051	.188	.846
6. Back strength	*.797*	*.031*	*.119*	*.806*
	.862	—.056	—.145	.876
7. Leg strength	*.873*	*—.203*	*—.033*	*.896*
	.620	.308	.257	.738
8. Shoulder strength	*.627*	*.304*	*.187*	*.722*
	.760	.167	.142	.791
9. Grip strength	*.767*	*.198*	*—.147*	*.806*
	1.019	.000	.000	1.019
10. Total strength	*.995*	*—.020*	*—.051*	*.996*
	.330	.729	—.219	.829
11. Sargent Jump	*.345*	*.739*	*—.170*	*.833*
	.519	.000	.524	.738
12. Weight	*.629*	*—.152*	*.323*	*.723*
	.....	.....	.....	.....
13. Computed Speed factor	*—.004*	*.877*	*—.005*	*.877*

Figures in roman type, upper left of each space, are for the first factorial matrix. Figures in italics, lower right of each space, are for the second factorial matrix. The computed speed factor was computed from the first factorial matrix and hence was not included in it.

The major contributions of the study seem to be the following:

(1) It is found possible to isolate and measure in relatively pure form, and with a high degree of accuracy, a factor which appears in the variables studied only as a compound (in oblique relationship). This is done by computing a multiple correlation, with 50 assumed as the mean and 10 as the standard deviation for the factor.

(2) When replaced in the correlation matrix, this new variable (the speed factor) falls on the axis as a primary trait.

(3) The variable may be used, practically, to compute, indirectly, the potential movement speeds of the body parts for this type of motor activity.*

(4) This technique may be used to isolate other factors in this or in other factorial matrices. The factor loadings are used as the correlations of the independent variables with the factor, and the variables selected should consist of one or more that are heavily weighted with the factor to be measured, and one or more with zero loadings with that factor, but with relatively high correlations with the other variables used and with the other factors, in order that the influence of these other factors may be partialed out.

* It has been so used by one coach of track and field athletics, who found that all boys showing a large amount of this factor (T-scores 65 and up) made excellent track athletes when trained, regardless of their previous experience.

TABLE III

Correlation Matrix

College Men

	Variable Number	1	2	3	4	5	6	7	8	9	10	11
4-pound shot	1											
8-pound shot	2	.751										
12-pound shot	3	.711	.811									
16-pound shot	4	.785	.846	.911								
24-pound shot	5	.709	.806	.815	.858							
Baseball put	6	.613	.611	.545	.589	.474						
Strength index	7	.475	.539	.468	.429	.488	.209					
Athletic strength index	8	.488	.587	.509	.501	.483	.406	.860				
Sargent Jump	9	.577	.444	.333	.392	.247	.312	.249	.122			
Weight	10	.388	.542	.512	.523	.596	.229	.638	.645	.038		
Speed factor	11	.835	.587	.502	.585	.494	.465	—.043	.467	.628	.045	

TABLE IV

Correlation Matrix

Junior High School Girls

	Variable Number	1	2	3	4	5	6	7	8	9	10	11	12	13
Dash	1													
Standing broad jump	2	.592												
Basketball throw, distance	3	.556	.607											
3-pound shot	4	.541	.599	.725										
12-pound shot	5	.410	.463	.735	.693									
Back strength	6	.195	.285	.512	.513	.628								
Leg strength	7	.194	.249	.439	.473	.508	.623							
Shoulder strength	8	.299	.429	.581	.525	.629	.529	.397						
Grip strength	9	.419	.352	.538	.496	.617	.604	.522	.567					
Total strength	10	.287	.357	.589	.592	.684	.844	.908	.638	.741				
Sargent Jump	11	.595	.731	.563	.582	.531	.252	.215	.420	.390	.332			
Weight	12	.090	.020	.341	.330	.516	.489	.412	.461	.606	.558	.000		
Speed factor	13	.472	.778	.469	.592	.335	—.012	—.142	.239	.054	—.004	.613	—.095	

TABLE V

Centroid Factor Matrix
(Second factoring)

College Men

	I	II	III
4-pound shot	.882	.276	—.252
8-pound shot	.903	.080	.121
12-pound shot	.865	.109	.316
16-pound shot	.902	.200	.302
24-pound shot	.840	.048	.391
Baseball put	.623	.210	—.034
Strength index	.636	—.653	—.103
Athletic strength index	.730	—.569	—.294
Sargent Jump	.488	.295	—.018
Weight	.591	—.467	.306
Speed factor	.665	.508	—.433

TABLE VI

Centroid Factorial Matrix
(Second factoring)

Junior High School Girls

Dash	.584	.252	—.231
Standing broad jump	.695	.450	—.459
Basketball throw, distance	.824	.183	.131
3-pound shot	.823	.231	.036
12-pound shot	.834	.046	.269
Back strength	.703	—.363	.158
Leg strength	.636	—.615	.147
Shoulder strength	.708	—.035	.138
Grip strength	.741	—.284	—.144
Total strength	.828	—.554	.024
Sargent Jump	.663	.369	—.343
Weight	.483	—.368	—.394
Speed factor	.454	.708	—.247

PSYCHOMETRIKA—VOL. 5, NO. 3
SEPTEMBER, 1940

WEBER LAWS, THE WEBER LAW, AND PSYCHOPHYSICAL ANALYSIS

ALSTON S. HOUSEHOLDER AND GALE YOUNG
THE UNIVERSITY OF CHICAGO

In failing to define the units in which the stimulus is to be measured, the Weber law might seem to make no definite assertion, and indeed, it is shown that any single empirical function, supposed to relate a given stimulus intensity with that intensity which is just noticeably greater, can be put into the Weber form by a suitable change of scale in which the stimulus intensity is to be measured. Nevertheless, it turns out that if different individuals have different Weber functions, when the intensities are measured on a given scale, then it is by no means always possible to transform the scale so that all of the functions can take on the Weber form. Some necessary conditions are given for the possibility of such a transformation when there is at hand a finite number of functions, and when the functions depend upon a single parameter the necessary and sufficient condition is easily derived. The same discussion leads to a generalization of Thurstone's psychophysical scale and shows that such a scale is always possible.

In discussing the discrimination of intensities there are two empirical questions that are not always clearly distinguished in the literature. On the one hand, to a graded stimulus we may have a graded response, in the sense that if a stimulus of intensity S is applied a reaction of intensity R results, and the result of increasing the intensity of the stimulus is to increase the intensity of the response without otherwise altering it. Thus experimentally R is obtainable as a function of S which is presumably monotonically increasing and continuous. Then one could choose some fixed quantity ΔR and ask for that value of ΔS associated with each S such that the responses elicited by S and by $S + \Delta S$ differ in the amount ΔR. Then by keeping ΔR fixed but varying S one obtains ΔS as some function of S. A function of this type could be taken as giving the discriminating power of the organism, and frequently is so taken when studies are made of tropisms in lower organisms (3).

But in the study of the higher vertebrates, at least, a somewhat different problem is usually set. Here stimuli of different intensities may lead to responses that differ qualitatively rather than quantitatively—a rat can be trained to jump to the right on seeing a light of given brightness, and to jump to the left when the light is of a different brightness. Then the intensities are said to be "discriminated",

or the difference to be "noticeable" by the organism, provided the responses called forth are qualitatively different. Then one can ask this question: If a stimulus of given intensity S calls forth a given response, by what amount ΔS must the intensity be increased in order to elicit a qualitatively different response? (It is understood that that ΔS is to be taken for which the response-difference occurs in a suitable, preassigned percentage of trials.) Again ΔS is empirically obtainable as a function of S, but *a priori* no relation can be predicted between this function and the one defined above. In the literature the two functions are not always clearly distinguished, although the very possibility of obtaining qualitatively different responses to quantitatively graded stimuli implies the presence in the nervous system of some kind of a "straining" mechanism, not required for merely quantitatively graded responses, which has the property of diverting excitations of different intensities along different nervous pathways (see, e.g. 5, 6, 8, and 9). Perhaps the reason for the confusion can be traced back to Fechner, who supposed that there was some intermediate psychic process of measure R operating in discrimination of the second type, which is such that ΔR must exceed a certain fixed value before such discrimination can be made, i.e. before the overt response can undergo a qualitative change.

Whatever the reason may be, the two functions are distinct, and we are here interested in the function of the second type. Thus, for any sense modality, and for any given choice of physical units in which to measure the intensity S of the stimulus, there is an empirically determinable function ΔS of S such that a change in S in the amount of ΔS brings forth a qualitative change in the reactions on the part of the organism being investigated. Otherwise stated, for any intensity S_1 there is a minimal stronger stimulus S_2 whose value is a function of S_1

$$S_2 = f(S_1)\,, \tag{1}$$

for which the response changes qualitatively in a given percentage of the trials. Any such function $f(S)$ so determined we shall call *a Weber function,* and if, in particular, $f(S)$ is linear and homogeneous:

$$f(S) \equiv aS\,, \tag{2}$$

where a is a constant, then the function will be said to have *the Weber form,* or to be, simply, linear. Note that this constant a exceeds by unity the ordinary "Weber constant" $\Delta S/S$.

Weber's law asserts, as is well known, that the Weber function has the Weber form. But Weber's law fails to define the units in which the stimulus is to be measured, whereas obviously, if the Weber

function has the Weber form with one choice of units, it will not necessarily have this form when the units are changed. Thus suppose that for some choice of units S *the* Weber law is verified empirically. Then if one should introduce new units $\bar{S}$ defined by

$$\bar{S} = \log S\,, \quad S = e^{\bar{S}}$$

we have

$$\bar{S}_2 = \bar{a} + \bar{S}_1\,, \qquad \bar{a} = \log a\,,$$

so that the associated Weber function has no longer *the* Weber form. Thus with any arbitrary choice of units for the measurement of the intensity of the stimulus of a given kind, there will be associated a particular Weber function whose form is empirically determinable. If the function does not have the Weber form, it may be that by a new choice of units the associated Weber function would turn out to have this form. Hence we may ask the question: Given any Weber function, supposed to have been determined empirically after fixing upon a definite scale of units, what change of scale can be employed in order that the Weber function associated with the new scale units shall have the Weber form? If it should turn out that for an arbitrary function $f(S)$, supposed to be a Weber function, it is always possible to introduce a scale transformation which will give the transformed Weber function the Weber form, then the Weber law is not a law at all, and is of no value unless it be taken as a definition of the scale units or a class of scale units, viz. that scale in which the associated Weber function has the Weber form. But if it should turn out that functions $f(S)$ exist of the type that might *a priori* be Weber functions but such that they cannot be transformed to the Weber form by any choice of scale units, then the Weber law can be made definite if it be taken as affirming that *in fact* no empirical Weber function will ever be one of these functions $f(S)$. This statement could be verified or refuted empirically, at least if we may be permitted to disregard errors of observation.

This is the problem for consideration, and we now define it more precisely. Let x or y, instead of S, represent the measure of the stimulus intensity in any scale. Let

$$\bar{x} = f(x) \tag{3}$$

or

$$\bar{y} = g(y) \tag{4}$$

denote the measure of that stimulus which is judged just greater than x or y in a suitable, preassigned percentage of the trials. Evidently always

$$f(x) > x\,, \qquad g(y) > y\,, \tag{5}$$

at least if this "discriminal" percentage is taken large enough, as we shall suppose for the present. We suppose, moreover, that these functions are properly monotonically increasing, continuous functions of their respective arguments. Finally it is assumed that the "intensity" is an essentially positive quantity, so that arguments and functions assume only positive values, and that zero intensity is associated with the absence of the stimulus.

Evidently the Weber functions $f(x)$ or $g(y)$ can be defined over a finite range of values of x or y only: too heavy a weight cannot be lifted, too bright a light is injurious, etc. Hence on any scale there will be a maximal x or a maximal y for which $f(x)$ or $g(y)$ is defined. It may be that at this maximal value the corresponding inequality (5) must be replaced by an equality.

There will also be a minimal x or y, different from zero for which $f(x)$ or $g(y)$ is defined, because of the presence of a finite threshold. One could define $f(0)$ or $g(0)$ as the threshold value, which would be equivalent to including the absolute threshold among the differential thresholds, but this we shall not do. Hence $f(x)$, or $g(y)$, is defined for only those values of x or y that are supraliminal.

Now suppose a given scale is chosen, the x-scale, and that the associated Weber function $f(x)$ is determined empirically. Let $g(y)$ be a function of prescribed form, arbitrary except for the restrictions imposed above, and let us inquire whether there exists a scale transformation

$$y = \phi(x) \tag{6}$$

which is such that the function $f(x)$ transforms into $g(y)$ by this scale transformation. We suppose that $\phi(x)$ is essentially positive, and is properly monotonically increasing and continuous in x, and that it is defined over the range for which $f(x)$ is defined. Naturally its values, as x varies within this range, must all be included in the range of values of y over which $g(y)$ is defined.

If the function ϕ is to satisfy these requirements, then since we must have, for any x,

$$\bar{y} = \phi(\bar{x}) = \phi[f(x)]\,,$$

it follows that ϕ must be so chosen that

$$\phi[f(x)] = g[\phi(x)] \tag{7}$$

for every x for which $f(x)$ is defined. This condition is necessary and it is evidently also sufficient. We therefore inquire whether, for a

given $f(x)$ and a prescribed $g(y)$, a function $\phi(x)$ can be found satisfying (7).

If the range of definition of $g(y)$ is sufficiently great, such a function ϕ can always be found, and in infinitely many ways, as follows. Let x_0 be the absolute threshold intensity in the x-scale, and let y_0 be any positive number lying on the range of definition of $g(y)$. Then define a sequence x_i and a sequence y_i by means of the following recursion formulas:

$$x_{i+1} = f(x_i) , \qquad y_{i+1} = g(y_i) . \tag{8}$$

Both sequences are properly monotonically increasing because of conditions (5). If at the upper limit of the x-range inequality (5) still holds, there will be a finite value, $n - 1$, of the index i such that $f(x_{n-1})$ equals or exceeds this limit. Then x_n is defined to be this limit and the first of equations (8) is taken only for $i < n - 1$. Then if the range of definition of g is great enough, or if y_0 is small enough, the sequence y_i will include terms up to and including y_n.

If, on the other hand, at this upper limit for which $f(x)$ is defined the inequality (5) must be replaced by an equality, then there will be infinitely many terms in the sequence x_i, and the limit of this sequence is this upper limit, which we shall call x_∞. Because the continuity of $f(x)$ we have $f(x_\infty) = x_\infty$. In this case, if g satisfies (5) wherever it is defined, g must be defined for all positive values of y, and the terms of the sequence y_i become infinite. But if $g(y) = y$ for some value of y, then again the sequence y_i has infinitely many terms which approach this value of y as a limit, whatever our choice of y_0 may have been.

In either of the various cases, $\phi(x)$ can be defined arbitrarily as a properly monotonically increasing continuous function of x over the interval $x_0 \leq x \leq x_1$, satisfying

$$\phi(x_0) = y_0 , \qquad \phi(x_1) = y_1 . \tag{9}$$

Now if x is in the interval $x_0 \leq x \leq x_1$, and y is in the interval $y_0 \leq y \leq y_1$, then evidently $\bar{x}$ and $\bar{y}$ as defined by (3) and (4) satisfy $x_1 \leq x \leq x_2$, and $y_1 \leq \bar{y} \leq y_2$. Hence if ϕ is defined over the interval from x_1 to x_2 by means of the equation

$$\phi(\bar{x}) = \bar{y} \tag{10}$$

it is evident that (7) is satisfied for every x in the first interval from x_0 to x_1. We can now proceed sequentially to define ϕ over the remaining intervals, using (3), (4) and (10) to define ϕ over any

interval $x_i \leq \bar{x} \leq x_{i+1}$ after it has been defined over the preceding one, and equation (7) will be satisfied wherever it is meaningful.

We have therefore shown that if there is given a single Weber function $f(x)$ in terms of some scale units x, and if $g(y)$ is an arbitrary, preassigned function, both functions being properly monotonically increasing, continuous functions satisfying inequalities (5) except, perhaps, at the upper limits of the ranges over which they are defined, then it is always possible to determine, in infinitely many ways, a change of scale which transforms the Weber function $f(x)$ into the Weber function $g(y)$, provided only that the range over which $g(y)$ is defined is sufficiently large.

In particular, *a single Weber function can always be given the Weber form by a suitable change of scale.*

It is reasonable to expect, however, that different individuals will have different Weber functions, though evidently for a given modality one would not wish to use different scales for different individuals. Thus presumably for one individual would be found a function f_1 in the x-scale, for another f_2 in the same x-scale, and if any transformation ϕ is to be made to a y-scale, the same transformation would have to be made on both f_1 and f_2. Given then, any finite set of empirical Weber functions, $f_1, f_2, f_3, \cdots$, will there always exist a scale transformation which transforms all these functions simultaneously into the Weber form (or some other specified forms $g_1, g_2, g_3, \cdots$)?

Before discussing this problem let us note a possible modification of the procedure for calculating the ϕ which transforms a single f into a single g. Instead of requiring that x_0 be the absolute threshold value, we could have taken as x_0 any value larger than this. For values of x greater than x_0 the procedure would be as before. In order to determine ϕ for values of x less than x_0, we note that the equations (3) and (4) can always be solved uniquely for x and y as functions of $\bar{x}$ and $\bar{y}$ respectively, because of the monotonicity and continuity of the functions of f and g. Hence along with the increasing sequences x_i and y_i defined by (8) we may place the decreasing sequences x_{-i} and y_{-i} defined by

$$x_{-i} = f(x_{-i-1}), \qquad y_{-i} = g(y_{-i-1}). \tag{11}$$

Thus we have a sequence of intervals to the left of x_0. Now if

$$\bar{y} = \phi(\bar{x}), \qquad x_0 \leq \bar{x} \leq x_1$$

and if x and y are defined by (3) and (4), then $x_{-1} \leq x \leq x_0$ and $y_{-1} \leq y \leq y_0$, and if we define $y = \phi(x)$, condition (7) is again satisfied wherever it is meaningful. We may now proceed sequentially

over the intervals to the left until ϕ has been defined wherever f is defined.

Hereafter we consider only functions $g(y)$ of the form

$$g(y) = ay\,, \qquad a = \text{const.}, \tag{12}$$

and inquire only about transformations yielding the Weber form. Suppose we have two functions $f_1(x)$ and $f_2(x)$, and suppose that for some $x = x_0$ on the range common to both functions we have

$$f_1(x) > f_2(x)\,. \tag{13}$$

Let y_0 be any positive number, and let the point (x_0, y_0) be our starting point for defining ϕ. We have, now, two increasing x-sequences, x'_i and x''_i, defined by

$$x'_i = f_1(x_{i-1})\,, \qquad x''_i = f_2(x_{i-1})\,. \tag{14}$$

Suppose that $f_1(x)$ is to be transformed into $g_1(y) = a_1 y$, and $f_2(x)$ into $g_2(y) = a_2 y$. The two y-sequences are

$$y'_i = a_1 y'_{i-1}\,, \qquad y''_i = a_2 y''_{i-1}\,, \tag{15}$$

and ϕ must be so defined that for every i,

$$y'_i = \phi(x'_i)\,, \qquad y''_i = \phi(x''_i)\,.$$

Now from (13) it follows that $x'_i > x''_i$, and since ϕ is monotonic increasing we must have $y'_i > y''_i$. But this is true if and only if $a_1 > a_2$. From this it is evident that the transformation cannot be made unless inequality (13) holds for all values of x on the common range. For if at some point $x = \bar{x}_0$ we should have the inequality in (13) replaced by equality, then we could use x_0 as our starting point and conclude that $a_1 = a_2$.

If at some point of the range common to two empirical Weber functions $\mathrm{f}_1(\mathrm{x})$ *and* $\mathrm{f}_2(\mathrm{x})$ *we have inequality* (13) *holding, then for* f_1 *and* f_2 *to be transformable to the Weber forms* $\mathrm{a}_1\mathrm{y}$ *and* $\mathrm{a}_2\mathrm{y}$, *respectively, by the same scale transformation* $\phi(\mathrm{x})$, *it is necessary that inequality* (13) *hold over this entire common range, and that* $\mathrm{a}_1 > \mathrm{a}_2$. This condition is by no means sufficient, as we can easily see. Thus, suppose the condition is satisfied, and, for simplicity, suppose x_0 is the larger of the absolute thresholds. Then we have two sequences x'_i and x''_j. Because of (13) it is evident that not more than one x'_i falls at an endpoint of or within any interval (x''_j, x''_{j+1}). Hence from the sequence x''_j we can select a subsequence x''_{j_ν} such that

$$x''_{j_\nu} \leq x'_\nu < x''_{j_{\nu+1}}\,. \tag{16}$$

Corresponding to this subsequence is the subsequence $y''_{j_\nu} = a_2^{j_\nu} y_0$ of y''_j and evidently we must have

$$a_2^{j_\nu} \leq a_1^{\nu} < a_2^{j_{\nu+1}}, \tag{17}$$

since ϕ has to assume the values of the y'_ν when x has the value x'_ν, and also the value y''_j when x has the value x''_j. The above inequalities can be put into the simpler form

$$j_\nu \leq \nu\alpha < j_\nu + 1, \qquad \alpha = \frac{\log a_1}{\log a_2}, \tag{18}$$

and this must hold for every positive integer ν corresponding to which there is a term of the sequence x'_ν. The equality can occur in (18) only when the equality occurs in (16). If the equality occurs once, α is uniquely determined. If it occurs twice the two equations may or may not be consistent. And examples of sequences j_ν of integers can easily be found in which the inequalities (18) are not all consistent.

If inequalities (18) are all consistent, then we may proceed as follows. Solve for x in all equations of the form

$$x_j = f_1(x), \qquad x_j = f_2(x), \tag{19}$$

where x_j is a member of one of the original sequences, or is one of the solutions resulting from the successive solutions of such equations. If y_j is the value of y associated with x_j, then y_j/a_1 or y_j/a_2 is the value that must be associated with x. Proceeding thus we obtain additional points through which the graph of $\phi(x)$ must pass, and hence new inequalities in addition to (18) which a_1 and a_2, or α, must satisfy. If all inequalities obtained in this way are consistent, and an α is chosen to satisfy them, then it may be possible to define ϕ sequentially over intervals bounded by the points x and the points of the original sequences.

When n functions $f_1, f_2, \text{----}, f_n$ are at hand, corresponding to n individuals, one has initially n sequences $x_i^{(j)}$ from the points of which one obtains other points by successive solutions of equations such as (19). From these one obtains certain inequalities on the possible values of the a_i after which one may be able to define ϕ as before. Thus:

If $f_1, f_2, \text{----}, f_n$ *are in empirical Weber functions for the same modality, corresponding to* n *individuals, it is not in general possible to determine a scale transformation* $\phi(x)$ *in such a way that the transformed functions all have the Weber form. For this to be possible it is necessary but not sufficient, that, for every* i *and* j, *if* $f_i(x) > f_j(x)$ *for some* x, *then this inequality holds for every* x *on the range*

common to both functions. Further necessary conditions are obtained in the form of sets of inequalities by starting with a fixed x_0, *defining additional points in all possible ways by* $x_j = f(x_k)$ *where* f *is any one of the* n *functions, and either* x_j *or* x_k *is a point previously determined. Corresponding to each* x_j *so determined, the value* y_j *that* $\phi(x_j)$ *must have can be determined in terms of an arbitrary initial* y_0 *and the constants* a *entering in the Weber forms. The inequalities arise from the required monotonicity of the function* ϕ.

If f is supposed to have a particular analytical form as a function of x and a single variable parameter, b, then for an individual whose parameter is b, the stimulus intensity $\bar{x}$ which is just noticeably greater than x is given by

$$\bar{x} = f(b, x) . \tag{20}$$

In this case the question whether there exists a scale transformation $y = \phi(x)$ such that $f(b, x)$ takes the Weber form $g = ay$ can be given a very simple answer. For the equation

$$\bar{y} = ay \tag{21}$$

can be regarded as defining a transformation of the variable y into $\bar{y}$, and we have in (21) a one-parameter continuous group of transformations. By a change of variable the group (21) is transformed into the group of translations

$$\bar{z} = c + z \quad (z = \log y , \quad c = \log a). \tag{22}$$

Now every change of variable transforms a one-parameter continuous group into a one-parameter continuous group, and conversely, every one-parameter continuous group can be transformed, by a suitable change of variable, into the translation group (22) and hence into the "Weber" group (21) (see, e.g. 4, Vol. 2, Part 2, p. 89). Thus if $f(b, x)$ is to be capable of taking the Weber form it must define a continuous group:

The necessary and sufficient condition for a Weber function of the form (20), *with continuous partial derivatives of the first order, to be transformable into the Weber form is that* (20) *defines a one-parameter continuous group of transformations. The scale transformation is, in this case, unique.*

Consider, now, a single individual, or a group of individuals regarded as identical for purposes of the investigation, and suppose a large number of judgments are made comparing stimuli of different intensities. Suppose we do not fix a "discriminal" percentage, but inquire, for each pair of stimuli, x and $\bar{x}$, as to the probability distri-

bution of judgments "$\bar{x} > x$". For each $\bar{x}$ and for each x there will be a unique probability p of the judgment "$\bar{x} > x$", and conversely for each x and each p there will be an $\bar{x}$ such that the probability of the judgment "$\bar{x} > x$" is p. Thus we can express

$$\bar{x} = f(p, x), \tag{23}$$

where p is the probability of the judgment "$\bar{x} > x$", and $\bar{x}$ should be a properly monotonically increasing function of p for each x, and of x for each p. We shall suppose, further, that

$$0 = f(0, x), \qquad \infty = f(1, x) \tag{24}$$

$$x = f(\tfrac{1}{2}, x) \tag{25}$$

identically in x, and that, ruling out equality judgments,

$$x = f(1 - p, \bar{x}). \tag{26}$$

Since to every $\bar{x}$ and x there corresponds a p such that (23) is satisfied, it is evident that, if we regard f as defining a transformation, the resultant of two transformations of the form (23) is again a transformation of the form (23). Moreover, these transformations include the identity (25) and the inverse (26).

Hence, the function $f(p, x)$ defines a one-parameter continuous group and is therefore reducible to the Weber form.

Furthermore, since every one-parameter group is also reducible to the translation group, we can, by suitable change of variable and parameter,

$$y = \phi(x) \qquad c = \psi(p) \tag{27}$$

reduce (23) to the form

$$\bar{y} = y + c. \tag{28}$$

Now c is a function of the probability of the judgment "$\bar{x} > x$", or "$\bar{y} > y$", and c measures the "distance" $\bar{y} - y$ on the y-scale. This, however, is Thurstone's psychophysical scale (11) in a somewhat more general form. Whether we think of introducing new scale units for measuring the physical intensity, or for measuring a psychological "discriminal process" as some functions of the physical intensity, the results are formally identical.

Any probability distribution of judgments "$\bar{x} > x$" when expressed in the form (23) *defines a one-parameter continuous group of transformations, and from such a distribution one can always de-*

fine a psychophysical scale by (27) *and* (28), *the scale units being expressed by an appropriate function* $\psi(p)$ *of the probability* p *of the judgment* "$\bar{x} > x$".

So far as our acquaintance with the literature extends, the general question of the measure of the stimulus intensity is not ordinarily raised in discussing the "validity" of Weber's law. Boring (1), however, did raise this question, to which Johnson (7) replied that Fechner had clearly defined the stimulus as energy. Nevertheless, he admitted that this cannot be, in general, and suggested that power should be taken instead. But whether it be power or energy or any other, is this to be measured at the periphery, at the first-order neurons, or elsewhere?

Cobb (2) points out that the form of the Weber function varies with the setting in which the stimulus is placed and remarks that while for one setting the Weber law may be verified over a wide range, it will not necessarily be verified when the setting is changed. But from the interpretation of the "Weber law" given here, this conclusion is not obvious.

This work was aided by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

REFERENCES

1. Boring, E. G. Did Fechner measure sensation? *Psych. Rev.*, 1928, 35, 443-445.
2. Cobb, P. W. Weber's law and the Fechnerian muddle. *Psych. Rev.*, 1932, 39, 533-551.
3. Crozier, W. J. Déterminisme et variabilité dans le comportement des organismes. *Actualités Scientifiques et Industrielles*, 1935, No. 261. Paris: Hermann et Cie.
4. Goursat-Hedrick. A course in mathematical analysis. New York: Ginn and Co., 1917.
5. Householder, A. S. A neural mechanism for discrimination. *Psychometrika*, 1939, 4, 45-58.
6. Householder, A. S. A neural mechanism for discrimination. II. Discrimination of weights. *Bull. Math. Biophysics*, 1940, 2, 1-13.
7. Johnson, H. M. Did Fechner measure "introspectional" sensations? *Psych. Rev.*, 1929, 36, 257-284.
8. Landahl, H. D. A contribution to the mathematical biophysics of psychophysical discrimination. *Psychometrika*, 1938, 3, 107-125.
9. Landahl, H. D. A contribution to the mathematical biophysics of psychophysical discrimination. II. *Bull. Math. Biophysics*, 1939, 1, 159-176.
10. Thurstone, L. L. The law of comparative judgment. *Psych. Rev.*, 1927, 34, 273-286.
11. Thurstone, L. L. Psychophysical analysis. *Am. J. Psych.*, 1927, 38, 368-389.

PSYCHOMETRIKA—VOL. 5, NO. 3
SEPTEMBER, 1940

SCORING WEIGHTED MULTIPLE KEYED TESTS ON THE IBM COUNTING SORTER*

RALPH BEDELL
THE UNIVERSITY OF NEBRASKA

Tests or personal inventories with differential item response weights may be scored by means of punch card equipment. Detailed instructions are given for preparing the cards and scoring the forms. The scoring speed is approximately four to eight times that attained by manual scoring.

The method of scoring herein described for weighted multiple keyed tests by use of the International Business Machines Counting Sorter can be used by any clerk trained to do sorting and counting on the machine. The machine has a relatively low rental cost and is usually used for a wide variety of statistical work other than test scoring. This type of test scoring does not require modification of the machine nor in the administration of the tests. While the method may be adapted to any multiple keyed test such as an interest inventory or personality questionnaire, detailed explanation is here given for scoring the *Vocational Interest Blank for Women* by E. K. Strong. This test can be completely scored for all eighteen keys by one clerk at the rate of two tests per hour by using a machine operating at 220 cards per minute. Two clerks using one machine operating at 440 cards per minute can score for all eighteen keys and profile the results at the rate of four tests per hour. This scoring spee dis approximately that of other machine methods now being used, and four to eight times that attained by manual scoring. The scoring speed is not affected by the number of tests to be scored; consequently it is as useful in clinics scoring small numbers of tests at a time as in mass testing programs.

The method consists of punching one Hollerith card for each test item. The punches are coded to give the number of the item, and the weight for each possible response to that item in each key for the test. The cards are then sorted by hand according to the testee's responses to the test, and the various weights totaled for each key on the IBM counter. Totals are entered on a tabulation sheet on which the required simple computations are made.

* The author is indebted to H. M. Cox, University of Nebraska, for certain suggestions on the columnar coding and the method of checking herein presented.

Coding

To code the weights for each response to each item, it is necessary to provide for nine weights varying from — 4 through +4. Only four positions on a column are required to code the nine weights if combinations of +1, —1, +3, and —3 are used. These weight code equivalents with the Strong weight given first are: +1 = +1; +2 = —1, +3; +3 = +3; +4 = +1, +3; 0 = no punch; —1 = —1; —2 = +1, —3 ; —3 = —3; —4 = —1, —3. Since the card has twelve possible positions in each column, three keys can be coded in one column for one response. The three possible responses require three columns for these keys. Each succeeding three keys are coded in each succeeding three columns. Thus, for the eighteen keys, eighteen columns are needed, plus three columns for the number of the item.

To make it possible to construct a convenient record and tabulation sheet to be used by the operator, it is desirable to code the three keys in each column as follows:

	Position on Card	Weight Code
	1	+1
(Key A)	2	—1
	3	+3
	4	—3
	5	+1
(Key B)	6	—1
	7	+3
	8	—3
	9	+1
(Key C)	10 (0)	—1
	11 (x)	+3
	12 (y)	—3

The accompanying tabulation sheet (Figure II) shows the method of coding for the eighteen keys for this test. The first three columns are reserved for the number of the item. Column 4 carries the weights for the *L* response in the keys for Artist, Stenographer-Secretary, and English Teacher; column 5 carries the weights for the *I* responses in these keys; and column 6 the weights for the *D* responses. The system is illustrated by the coding for item *109 Stenographer L. I. D.* The Strong weights in each key for each response with the columns in which these weights are coded are given in Table I. The complete punching of the card for this item is given with the column

numbers in bold-face type and the position numbers in light face. **1**—1; **2**—0; **3**—9; **4**—2, 4, 5, 7, 0; **5**—2, 6; **6**—2, 3, 6, 8; **7**—6, 0; **8**—0; **9**—2, 5, 9; **10**—2, 6; **11**—1, 5; **12**—no punch; **13**—1, 4, 5, 8, x; **14**—2; **15**—2, 3, 5, y; **16**—5, 0; **17**—5, 0; **18**—5, 8, 0, x; **19**—2, 6, 0; **20**—1, 5; **21**—9. This punched card is shown in Figure I.

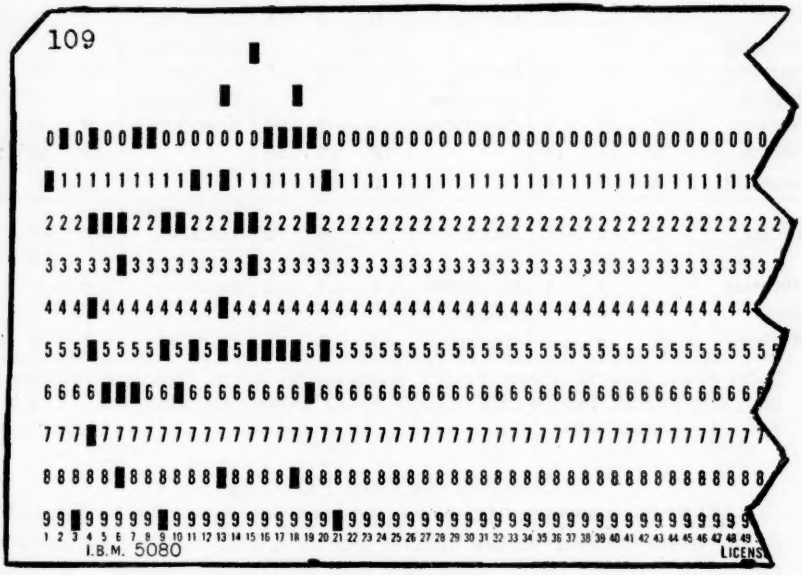

FIGURE I

The completely punched card for item 109. The position next above 0 is x and the top position is y.

One card is punched for each item making a deck of 410 cards. This deck is then interpreted* on the first three columns so that the number of each item appears on the card above columns 1, 2, and 3. The cards are then duplicated* and verified by the Reproducing Punch* and interpreted* to make as many decks as are needed. To avoid possible loss and excessive wear the deck of master cards is kept in a safe, and removed only to make the duplicate sets which are used for scoring.

* These are card-making operations, not test-scoring operations. The Interpreter is a machine which automatically prints at the top of each card information punched in the card. The Reproducing Punch is another machine which automatically duplicates and verifies new decks from the original punched deck. It is not necessary to lease these machines as the service here indicated may be obtained through any International Business Machines office at nominal charge.

TABLE I

Original Strong Weights in Each Key for Each Response to Item 109 with the Columns in Which the Weights Are Coded

Column	4	5	6		7	8	9		10	11	12
Response	*L*	*I*	*D*		*L*	*I*	*D*		*L*	*I*	*D*
Artist	—4	—1	+2	Lawyer	0	0	—1	Teacher in General	—1	+1	0
Stenog'er-Secretary	+4	—1	—4	Dentist	—1	0	+1	Math.-Phys. Sci. Tchr.	—1	+1	0
English Teacher	—1	0	0	Insurance Saleswmn.	—1	—1	+1	Femininity-Masculinity	0	0	0
Column	13	14	15		16	17	18		19	20	21
Response	*L*	*I*	*D*		*L*	*I*	*D*		*L*	*I*	*D*
Author	—2	—1	+2	Nurse	0	0	0	Soc. Science Teacher	—1	+1	0
Librarian	—2	0	+1	Housewife	+1	+1	—2	Y. W. C. A. Secretary	—1	+1	0
Office Worker	+3	0	—3	Social Worker	—1	—1	+2	Physician	—1	0	+1

Table reads: Code in column 4 the *L* responses for Artist, Stenographer-Secretary, and English Teacher; the original Strong weights for which are —4, +4, and —1 respectively.

Checking the Card Punching

While the accuracy of the original card punching may be checked in the usual manner on the verifier punch, a trial balance technique for using the counter to check accuracy has been developed. A balance sheet is set up which carries the algebraic totals of the weights provided by Strong for each twelve items for each response for each key. This gives thirty-six totals for each twelve items. These totals are then cumulated for each sixty items. The cards are then run through the counter in groups of sixty for each response for each key. If totals do not check, an error is indicated and that group of sixty is rerun through the machine in groups of twelve to check for these totals. The group of twelve which contains the error is then read visually for the key and the response on which the error is indicated.

Scoring Procedures

The scoring procedures are given in detail for one clerk using a machine running at 220 cards per minute.

1. Record on the tabulation sheet (Figure II) the name of the person whose test is being scored.
2. Hand sort cards into three stacks corresponding to the tes-

FIGURE II

Tabulation Sheet

Name .....

Directions: Record total for each counter opposite its position number and under the columns on which the count was made.

S_1 equals the algebraic sum of the numbers opposite 3 and 4, or 7 and 8, or 11 and 12.

$3S_1$ equals three times S_1.

S_2 equals the algebraic sum of the numbers opposite 1 and 2, or 5 and 6, or 9 and 10.

Tot. equals the algebraic sum of $3S_1$ and S_2.

Record the correct signs for S_1, $3S_1$, S_2, and Tot.

Position	Col. 4, 5, 6		Col. 7, 8, 9		Col. 10, 11, 12		Col. 13, 14, 15		Col. 16, 17, 18		Col. 19, 20, 21	
1	+		+		+		+		+		+	
2	−		−		−		−		−		−	
3	+		+		+		+		+		+	
4	−		−		−		−		−		−	
S_1												
$3S_1$												
S_2												
Tot.												
Key	Artist		Lawyer		Teacher in Gen.		Author		Nurse		Soc. Sci. Teacher	
5	+		+		+		+		+		+	
6	−		−		−		−		−		−	
7	+		+		+		+		+		+	
8	−		−		−		−		−		−	
S_1												
$3S_1$												
S_2												
Tot.												
Key	Stenog.-Sec.		Dentist		Math. Teacher		Librarian		Housewife		YWCA Sec.	
9	+		+		+		+		+		+	
10	−		−		−		−		−		−	
11	+		+		+		+		+		+	
12	−		−		−		−		−		−	
S_1												
$3S_1$												
S_2												
Tot.												
Key	English Teacher		Insurance Saleswmn.		Femin.-Masc.		Office Worker		Social Worker		Physician	

This Tabulation Sheet is expanded and mimeographed to fill an 8½ x 11 sheet.

tee's responses. This places all *L* responses in the first stack, all *I* responses in the second stack, and all *D* responses in the third stack.

3. Count the *L* stack on column 4 and remove this stack from the machine, keeping it separate from the other two stacks. Without resetting the counters, count the *I* stack on column 5 and remove this stack as above. Count the *D* stack on column 6 in the same manner.
4. Record the counter totals on the tabulation sheet.
5. Reset counters to zero.
6. Repeat 3, 4, and 5, except count the *L* stack on column 7,

the *I* stack on column 8, and the *D* stack on column 9. Continue this procedure, making sure that totals are recorded and counters reset to zero after the counts are completed on columns 6, 9, 12, 15, 18, and 21.

7. While the machine is making the above counts, perform the computations indicated on the tabulation sheet.
8. After all counts are completed, place all cards in the machine and resort to the original numerical order by sorting on columns 3, 2, and 1.
9. While performing operation 8, note and hand sort the second deck as indicated in 1 and 2.

For a machine operating at 440 cards per minute, two clerks are necessary to keep the machine running with minimum time loss. One clerk operates the machine and records the totals. The other clerk does the hand-sorting and computations indicated in 1, 2, and 7 above, and in addition profiles the results or writes them on the test blanks according to the system desired.

Some Precautions

To save time due to wear on the cards, it is desirable to have available five decks for each machine. One of these decks is set aside to be "robbed" of individual cards that may be unduly roughed or torn in the machine. Replacing such worn cards in the deck being used is much more economical than trying to force worn cards through the machine or stopping to reproduce the cards on the Alphabetic Duplicating Punch. Such removed worn cards can be set aside, later duplicated, and the new cards returned to the "robbed" deck. As clerks sometimes find time to keep their sorting ahead of the machine, the remaining four decks are used for scoring. Clerks should handle all cards by the ends as handling on the sides increases machine wear. New cards should be examined to make sure they have the proper bend for the machine before being put into use. No errors made by the IBM Reproducing Punch have been found, and it is believed that checks of card punching beyond those indicated for the original deck are unnecessary. The accuracy of the counters on the machine can be checked by counting all of the cards on any one column, the totals for which have been determined previously when the machine was known to be accurate.

Adaptability to Other Multiple Keyed Tests

While the method has been given in detail for only the *Strong Vocational Interest Blank for Women*, adaptation can readily be made to several other tests. The *Strong Vocational Interest Blank for Men,*

the *Humm-Wadsworth Temperament Scale*, and the *Bernreuter Personality Inventory* illustrate the types of tests for which this scoring method is appropriate. For any test which has the same range of weights as the Strong blanks, it is believed that the above coding is the most economical that can be derived for this method of scoring. Tests that have a different range of weights will require different codings, but the fundamental principle remains the same. The scoring speeds mentioned at the beginning of this article can be compared with the speeds of other methods being used to determine in advance whether or not it would be economical to use the method here proposed. By checking the time clerks require for the hand-sorting, computing the time required by the machine, and estimating the extent to which operations may be carried on simultaneously, it is possible to arrive at a reasonably accurate estimate of the scoring speed that may be attained for any test.

PSYCHOMETRIKA—VOL. 5, NO. 3
SEPTEMBER, 1940

CONTRIBUTIONS TO THE MATHEMATICAL THEORY OF HUMAN RELATIONS III

N. RASHEVSKY

THE UNIVERSITY OF CHICAGO

In continuation of a previous study, somewhat more complex relations are considered for the case of interaction of two classes, of which one "organizes" the activities of the other. By way of illustration it is shown how different assumptions concerning the attitudes of the individuals to such an interaction of classes lead to different mathematical expressions.

In several previous publications (1, 2, 3, 4) we have attempted to describe in mathematical terms some social phenomena. As has been emphasized in those papers, such a description is at present possible only at the expense of a very great oversimplification and idealization of the phenomena to be described. As in every theoretical science in its early stages, we must confine ourselves to rather simple conceptual situations, which may bear only a very remote resemblance to actually observed phenomena. We hope eventually to come to the study of sufficiently complex situations, so that our formulae may be applied to real cases. But it is impossible to arrive at this result without a previous purely abstract study, that is without first studying the theory for theory's own sake. A more thorough discussion of this assertion is to be found in the above mentioned previous publications, as well as in other papers, to which references are given in the above mentioned ones.

The present paper is a continuation of the previous ones, and therefore presupposes a knowledge of these. In particular we shall have to refer to results obtained in the last two papers (3, 4), which, for abbreviation, will be referred to as HR I and HR II respectively.

In HR II we have investigated the interreaction of two types of individuals, who are characterized by the following purely hypothetical properties:

Each individual of type I produces per unit of time an amount p_1 of goods necessary for the maintenance of his life, and he consumes per unit time an amount c_1 of these goods. Similarly every individual of type II produces and consumes correspondingly amounts p_2 and c_2. Denoting by N_1 and N_2 the numbers of individuals of type I and

II respectively let us consider the case, in which

$$N_1 << N_2 \; ; \quad p_1 - c_1 = \varepsilon_1 > 0 \; ; \quad p_2 - c_2 = \varepsilon_2 < 0 \, . \tag{1}$$

Then in the absence of any exchange or external supply, the individuals of type I are capable of indefinite existence, those of type II are not.

Now let the individuals of type I be capable of *organizing* their work, as well as the work of individuals of type II in such a way that the production of necessary goods is largely increased. In other words let the production of goods require a certain amount of *organization skill*, which is possessed by individuals of type I but not by those of type II, so that when left to themselves the individuals of type II cannot produce enough goods, but they can do so under the direction of the individuals of type I.

In HR II we have investigated the following relation between the individuals of type I and II: The latter agree to work under the direction of the former, so as to increase their output of goods, provided they do receive a part of the goods thus produced for their consumption. On the other hand, the individuals of type I will agree to direct the work of the individuals of type II only if they receive also part of the goods thus produced. For simplicity's sake we have considered the situation, in which every individual of type II, who expends an amount of labor w_2 per unit time, gets a fraction θw_2 of goods per unit time. The quantity θ may be considered as the price of his labor, paid not in money, but in goods produced by himself.

We shall consider in this paper a somewhat more complex and perhaps more natural situation, in which an individual of type II receives goods not only in proportion to the amount of labor he expends per unit time, but also in proportion to the amount of goods which he produces per unit time under the direction of individuals of type I.

In HR II we considered that the total amount of goods produced per unit time by all individuals of type II under the direction of the individuals of type I is equal to (HR II, equ. 2):

$$P = AN_2 w_2 \, , \tag{2}$$

where A is a function of the ration $\eta = N_1/N_2$, $A = f(\eta)$, and $f(\eta)$ has the following properties:

$$f(\eta) \; > 0 \text{ for } \eta > 0 \; ; \quad f(0) = 0 \, , \quad f(\infty) = \text{const.} \tag{3}$$

The reasons for those properties have been explained in HR II.

Expression (2) tacitly implies, that the amount of supervising work done by individuals of type I per unit time is constant. It is of

interest to consider the more general case, in which P is proportional to the product $w_1 w_2$ of the amounts of labor w_1 and w_2 given per unit time by individuals of type I and II. P by definition is the amount of goods produced per unit time by individuals of type II *under the direction of individuals of type I*, or in other words P is the excess over the "natural production" $N_2 p_2$ of the individuals of type II when left to themselves. Therefore P should vanish for either $w_1 = 0$ or $w_2 = 0$. Thus we shall now consider the case, where

$$P = f(\frac{N_1}{N_2})\, N_2\, w_1 w_2 \,, \tag{4}$$

$f(\eta)$ being the same as before.

By an argument similar to that used in HR II, we now find, instead of equations (10) and (11) of HR II, the following equations for the rates of change of the total amount of goods W_1 and W_2 possessed by all individuals of type I and II correspondingly:

$$\frac{dW_1}{dt} = N_1 \varepsilon_1 + N_2 f(\eta)\; w_1 w_2 (1 - \theta) \,, \tag{5}$$

$$\frac{dW_2}{dt} = N_2 \varepsilon_2 + \theta N_2 f(\eta) w_1 w_2 \,. \tag{6}$$

From the definition of θ we always have $\theta < 1$.

In HR II we have supposed that the "price" θ is related to the amount of labor w_2 by a simple hypothetical demand equation:

$$w_2 = w_{02} - \frac{b_2}{\theta} \,, \tag{7}$$

where w_{02} and b_2 are constants. A similar situation may be considered for w_1. Individuals of type II will be willing to give more supervising labor, the greater the amount of additional goods they receive in return. We shall however introduce here in our assumptions an explicit asymmetry in the behaviour of the two classes. This is partly made in order to simplify some calculations. The effects of *not introducing* such an asymmetry will be discussed at the end of this paper

We shall assume that while the individuals of class II determine the amount of labor they are willing to give by the *fraction* of the goods which they receive in return per unit of goods produced, the individuals of class I determine the amount of supervising labor which they are willing to give by the *total* amount of goods, which they receive from class II per individual of class I as the result of cooperation. Thus we shall put

$$w_1 = w_{01} - \frac{b_1}{P(1-\theta)/N_1} .$$

Introducing for P its expression (4), we find:

$$w_1 = w_{01} - \frac{b_1\eta}{f(\eta)\, w_1 w_2 (1-\theta)} . \tag{8}$$

Equation (8) may be written:

$$f(\eta)\,(1-\theta)\, w_2 w_1^2 - w_{01} f(\eta)\,(1-\theta)\, w_2 w_1 + b_1\eta = 0 . \tag{9}$$

Solved with respect to W_1, it gives:

$$w_1 = \frac{w_{01}}{2} \pm \frac{\sqrt{w_{01}^2 f^2(\eta)\,(1-\theta)^2\, w_2^2 - 4b_1\eta f(\eta)\,(1-\theta)\, w_2}}{2 f(\eta)\,(1-\theta)\, w_2} . \tag{10}$$

We shall now discuss the second term on the right side of equation (5). This term represents the amount of goods received per unit time by all individuals of type I from those of type II. To this end consider the expression $w_1 w_2 (1-\theta)$. From equation (10) and (7) we find after rearrangements:

$$w_1 w_2 (1-\theta) = \frac{w_{01}}{2}(1-\theta)\,(w_{02} - \frac{b_2}{2})$$

$$\times \left[1 \pm \sqrt{1 - \frac{4b_1\eta}{f(\eta)\,(1-\theta)\, w_{01}^2 (w_{02} - b_2/\theta)}}\ \right] . \tag{11}$$

The expression

$$x = (1-\theta)\,(w_{02} - b_2/\theta) , \tag{12}$$

which occurs in the denominator under the radical sign has a maximum for

$$\theta_m = \sqrt{\frac{b_2}{w_{02}}} . \tag{13}$$

The maximum value x_{max} of x is equal to:

$$x_{max} = w_{02} + b_2 - 2\sqrt{b_2 w_{02}} . \tag{14}$$

If

$$\frac{4b_1\eta}{f(\eta)\, w_{01}^2 x_{max}} < 1 ,$$

or

$$x_{max} > \frac{4b_1\eta}{f(\eta)\, w_{01}^2} , \tag{15}$$

then the expression $w_1 w_2 (1 - \theta)$ has two real positive values for a given θ, provided θ lies within an interval

$$\theta_1 < \theta < \theta_2 < 1 . \tag{16}$$

which is in the neighborhood of the value given by equation (13); θ_1 and θ_2 are roots of the equation obtained by equating the right hand side of equation (12) to the right hand side of inequality (15).

If inequality (15) does not hold, then $w_1 w_2 (1 - \theta)$, and therefore the last term of equation (5) is imaginary. Under these conditions a cooperation of the two groups is impossible: no agreement as to price can be reached. It is interesting to note, that inequality (15), which represents the conditions for the possibility of cooperation, involve such quantities as the function $f(\eta)$, or the constant w_{01}, which are more or less of a physical nature, as well as the constant b_1, which is more of a psychological nature.

The existence of two values of the expression $w_1 w_2 (1 - \theta)$ is due to a two-valued nature of w_1, as defined by equation (8). This

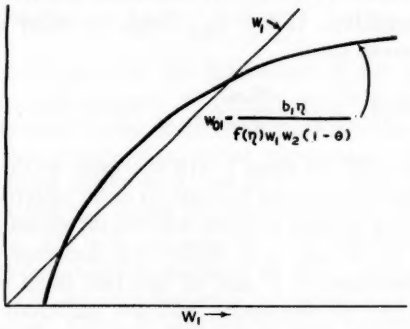

FIGURE 1

however is not connected with the specific analytical expression chosen for w_1, but lies rather in the nature of the general assumption made, namely that the relation of w_1 to the return in goods is such, that for returns below a certain value the individual will not wish to give any labor, while for increasing return the amount of labor invested increases, tending to a limiting value determined by the physically possible maximum of the labor given. This, combined with the assumption that the return in goods is in its turn proportional to the amount of labor given, leads to the two valued character of w_1 in terms of θ. For the left hand side of equation (8) is represented by a straight line (Figure 1), while the right hand side is represented

by a curve, which, regardless of its detailed analytic specification, *in general*, intersects the straight line in two points.

Returning now to equation (11), we see that the value of $w_1w_2(1 - \theta)$ which corresponds to the upper sign before the radical is always the largest of the two. For $\theta = \theta_m$ as given by equation (13), both the expression in brackets and the expression in front of it have a maximum. Hence for $\theta = \theta_m$, the quantity $w_1w_2(1 - \theta)$ has a maximum for the upper sign before the radical. For the lower sign it may or may not have a maximum.

In HR I we have established some equations describing the conditions under which one of the two classes controls the behaviour of the other regardless of such interreactions as studied here. If the relations between the constants, involving social control (HR I), are such, that in the present case the class I, which is composed of individuals of type I, controls behaviour of class II, (which is not *a priori* necessary), then class I will prescribe as price for the labor of class II the value θ_m, for this makes dW_1/dt a maximum.

Due to inequalities (1), dW_1/dt is always positive for $\theta < 1$. But dW_2/dt may be negative. If $\theta = \theta_m$, then, in order to make dW_2/dt also positive, we must have

$$\theta w_1 w_2 > -\frac{N_2 \varepsilon_2}{N_2 f(\eta)}, \qquad \text{for } \theta = \theta_m . \tag{17}$$

Otherwise the "wealth" of class I will increase while that of class II will decrease. Class II will not be able to exist indefinitely, and without it class I will not be able to have additional goods also. For a continued negativity of dW_2/dt will result in a decrease of N_2, and this will result in a decrease of P and of the last term of equation (5). Thus condition (16) is necessary, but not sufficient for the coexistence of such two classes. It must be supplemented by condition (17).

If class I fixes θ at θ_m, this will result in a maximum of dW_1/dt. But this does not mean that the rate of increase of the total goods for both classes, that is $d(W_1 + W_2)/dt$ will be a maximum. In fact, this will not be the case. From equations (5) and (6) it follows that $d(W_1 + W_2)/dt$ has a maximum for such a value θ'_m of θ, for which w_1w_2 has a maximum. Denote

$$y_1 = w_1w_2 > 0 \; ; \quad y_2 = w_1w_2(1 - \theta) = y_1(1 - \theta) > 0 . \tag{18}$$

The quantity θ'_m is defined as that value of θ which makes y_1 a maximum. Hence

$$\frac{dy_1}{d\theta} = \theta \text{ for } \theta = \theta'_m . \tag{19}$$

$$\frac{dy_1}{d\theta} = \theta \quad \text{for} \quad \theta = \theta'_m \,. \tag{19}$$

If θ_m is the value, which makes $y_2 = y_1(1 - \theta)$ a maximum, then

$$\frac{dy_2}{d\theta} = 0 \quad \text{for} \quad \theta = \theta_m \,.$$

But for $\theta = \theta'_m$, because of (19)

$$\frac{dy_2}{d\theta} = -y_1(\theta_m) + (1 - \theta_m)\frac{dy_1}{d\theta} = -y_1 < 0 \,. \tag{20}$$

Hence $\theta'_m > \theta_m$.

In other words, under such conditions as are described by equations (5), (6), (7) and (8), such a cooperation of the two classes, which is optimal for the population as a whole will not be optimal for class I.

Denoting by θ''_m the value of θ, which makes $\theta w_1 w_2$ a maximum, and therefore gives a maximum dW_2/dt, we can prove by a similar argument that

$$\theta''_m > \theta'_m > \theta_m \,. \tag{21}$$

If the class II controls the behaviour of the population, or if a third class is present, which, while not participating directly in the cooperation discussed here, influences nevertheless the behaviour of the population by virtue of a social mechanism discussed in HR I, then θ may be fixed either at θ'_m or at θ''_m. This will lead to optimal conditions for either the whole population or for class II. It may however happen, that $\theta''_m > \theta_2$, or even $\theta'_m > \theta_2$, [Equation (16)]. In this case no cooperation will be possible at all. Class I will continue to exist due to $\varepsilon_1 > 0$, though the rate of increase of its wealth will be very much reduced. Class II will not be able to exist indefinitely. Under the conditions studied here, the coexistence of the two classes may be possible, for a certain choice of constants, only when it corresponds to conditions close to those optimal for class I.

If we do not introduce the asymmetry, expressed in the different structure of equations (7) and (8), but take for w_1 an expression similar to (7), except that θ would be replaced by $(1 - \theta)$, we would have simplified the situation somewhat. We now would have:

$$(1 - \theta) w_1 w_2 = (1 - \theta)\left(w_{01} - \frac{b_1}{1 - \theta}\right)\left(w_{02} - \frac{b_2}{\theta}\right) \,, \tag{22}$$

which has a maximum for

$$\theta = \theta_m = \sqrt{\frac{w_{01} w_{02}}{(w_{01} - b_1) b_2}} \,. \tag{23}$$

Similarly as before we shall find $\theta''_m > \theta'_m > \theta_m$.

On the other hand, if we take for both w_1 and w_2 expressions of the form of equation (8), then the situation is mathematically more complex. Though essentially similar results are obtained, yet closed expressions are very cumbersome, for we have to deal now with two simultaneous quadratic equations, which reduce to a single fourth degree equation.

The different expressions that may be used for w_1 and w_2 correspond to different assumptions about the *psychological* attitudes taken by each class towards cooperation of the type described. The relations discussed here are too simple to be of any practical use. They merely *illustrate* how different psychological attitudes of different social classes may be in principle translated into mathematical language. But it is only through the study of such simple *illustrations* that we may hope to arrive at the study of more concrete real cases.

The author is indebted to Dr. A. S. Householder for a discussion of the paper and valuable critical comments.

REFERENCES

1. N. Rashevsky. Outline of a Mathematical Theory of Human Relations. *Philosophy of Science*, 2, 413-429, 1935.
2. N. Rashevsky. Further Studies on the Mathematical Theory of Human Relations, *Psychometrika*, 1, 21-36, 1936.
3. N. Rashevsky. Studies in Mathematical Theory of Human Relations, *Psychometrika*, 4, 221-239, 1939.
4. N. Rashevsky, 1939 b. Studies in Mathematical Theory of Human Relations II, *Psychometrika*, 4, 283-299, 1939.

PSYCHOMETRIKA—VOL. 5, NO. 3
SEPTEMBER, 1940

THE EVALUATION OF MULTIPLE AND PARTIAL CORRELATION COEFFICIENTS FROM THE FACTORIAL MATRIX*

P. S. DWYER
THE UNIVERSITY OF MICHIGAN

This paper shows how to compute multiple correlation coefficients, partial correlation coefficients, and regression coefficients from the factorial matrix. Special emphasis is given to computation technique and to approximation formulas. The method is extremely flexible in application since it may be applied to any subset of the original set of observed variables. It is also extremely useful when many of these coefficients are desired.

PART I

MULTIPLE CORRELATION COEFFICIENTS FROM THE FACTORIAL MATRIX

1. *Introduction.* One of the disadvantages of multiple correlation theory is the enormous amount of time needed to compute the multiple correlations when n is large. Even the Doolittle method, (1, 2, 3) which is being recognized as an efficient method of computing multiple correlations, demands a large amount of computation when n is large. It is shown in this part how multiple correlations can be obtained from the factorial matrix, and in particular, from a centroid solution, though any solution which gives communality and uniqueness factors is satisfactory. Though a centroid solution does take time, the total time necessary to obtain the multiple correlations is but a small fraction of that demanded by other methods if the number of variables is large, if the number of factors is small, and if a number of such coefficients is desired.

2. *Notation.* (a) *From multiple correlation theory.* We assume that there are n standard variables, $x_1, x_2, \cdots, x_i, \cdots, x_j, \cdots, x_n$.

*Editor's Note: The reader will recognize the subject-matter treated in this article as closely allied to that of an article by Louis Guttman, "Multiple Rectilinear Prediction and the Resolution into Components," in the June, 1940 issue of this journal. Although there is considerable overlapping in the topics considered in these two articles, it was felt that some readers would be more interested in one approach, while others would gain more from the other approach and that still others would find both presentations of value. Guttman's article is somewhat more concerned with the theoretical aspects of the problem, while Dwyer's article emphasizes the technique of computation and approximation. The manuscript for Dwyer's article was received while Guttman's article was in press. At Dwyer's request, Guttman made available to him a pre-publication copy of his manuscript so that Dwyer was able to insert in his manuscript references to the earlier one.

The variables are recorded for a population of N individuals and the matrix of the correlations, $[r_{ij}]$, is computed The determinant of this matrix is indicated by Δ and the determinant of the matrix formed by deleting variable x_j is denoted by Δ_j. If R_j indicates the multiple correlation of x_j with the $n - 1$ remaining variables, then by the usual formula (4, p. 301)

$$R_j = \sqrt{1 - \frac{\Delta}{\Delta_j}}. \tag{1}$$

$R_{j)i(}$ is used to indicate the multiple correlation of x_j with all the remaining variables except x_i. Similarly $R_{j)ik(}$ indicates the multiple correlation of x_j with all the remaining variables except x_i and x_k.

(b). *From multiple factor theory.* We assume that the correlation matrix has been subjected to an analysis which results in $r + n$ standard orthogonal variables, r of which are common variables and n of which are uniqueness variables. More specifically we assume that every x_j can be represented by

$$x_j = a_{j1}y_1 + a_{j2}y_2 + \cdots + a_{jk}y_k + \cdots + a_{jr}y_r + u_jz_j, \tag{2}$$

where $y_1, y_2, \cdots, y_r, z_1, z_2, \cdots, z_n$ are the orthogonal variables, and the values a_{jk} are known.

It follows at once, since $r_{ij} = \frac{\Sigma x_i x_j}{N}$, that

$$r_{ij} = a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{ik}a_{jk} + \cdots + a_{ir}a_{jr}, \quad i \neq j \tag{3}$$

and

$$1 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jk}^2 + \cdots + a_{jr}^2 + u_j^2, \tag{4}$$

as indicated by Thurstone (5, p. 66). Furthermore we define $h_j^2 = a_{j1}^2 + a_{j2}^2 + \cdots + a_{jr}^2$ to be the communality of variable j and u_j^2 to be the "uniqueness." Then

$$h_j^2 + u_j^2 = 1 \quad \text{and} \quad u_j = \sqrt{1 - h_j^2}. \tag{5}$$

We next define the common part of x_j to be that part of x_j which is expressible in terms of the common variables y_k. Then

$$x'_j = a_{j1}y_1 + a_{j2}y_2 + \cdots + a_{jk}y_k + \cdots + a_{jr}y_r.$$

It is to be noted that x'_j is not a standard variable since its variance is not unity. However, if every factor loading a_{jk} is divided by h_j we get the standard variable

$$x''_j = \sum_{k=1}^{r} A_{jk}y_k,$$

where $A_{jk} = \frac{a_{jk}}{h_j}$. The factor loadings are increased by this process and hence we use the term "augmented" (5, p. 158). Furthermore, we introduce the term "communality augmented" to distinguish these from the "uniqueness augmented" loadings introduced in this paper.

We let $U^2 = [u_i^2]$ indicate the diagonal matrix composed of the uniqueness loadings. Then (3), (4), and (5) can be combined to give the fundamental matrix equation

$$R = AA' + U^2 . \tag{6}$$

3. *The value of* R_j. R_j can be expressed in terms of the factor loadings if Δ and Δ_j can be so expressed. Now Δ_j can be so expressed if Δ can. Hence such an expression for Δ will, essentially, solve our problem. A general development using matrix terminology is preceded by a special case, $n = 3$ and $r = 2$, in which determinant theory is used. We first express Δ in the form

$$\Delta = \begin{vmatrix} 1 & r_{12} & r_{13} & a_{11} & a_{12} \\ r_{21} & 1 & r_{23} & a_{21} & a_{22} \\ r_{31} & r_{32} & 1 & a_{31} & a_{32} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} \tag{7}$$

and eliminate the correlation coefficients by multiplying column 4 by $-a_{11}$, column 5 by $-a_{12}$, and adding to column 1; by multiplying column 4 by $-a_{21}$, column 5 by $-a_{22}$, and adding to column 2; by multiplying column 4 by $-a_{31}$, column 5 by $-a_{32}$, and adding to column 3. We then have

$$\Delta = \begin{vmatrix} u_1^2 & 0 & 0 & a_{11} & a_{12} \\ 0 & u_2^2 & 0 & a_{21} & a_{22} \\ 0 & 0 & u_3^2 & a_{31} & a_{32} \\ -a_{11} & -a_{21} & -a_{31} & 1 & 0 \\ -a_{12} & -a_{22} & -a_{32} & 0 & 1 \end{vmatrix} . \tag{8}$$

We now divide the first row and the first column by u_1, the second row and the second column by u_2, the third row and the third column by u_3, and we get, if we let $\frac{a_{jk}}{u_j} = B_{jk}$, the "uniqueness augmented" loading,

$$\Delta = u_1^2 u_2^2 u_3^2 \begin{vmatrix} 1 & 0 & 0 & B_{11} & B_{12} \\ 0 & 1 & 0 & B_{21} & B_{22} \\ 0 & 0 & 1 & B_{31} & B_{32} \\ -B_{11} & -B_{21} & -B_{31} & 1 & 0 \\ -B_{12} & -B_{22} & -B_{32} & 0 & 1 \end{vmatrix} . \tag{9}$$

This can be further simplified by multiplying the first row by B_{11}, the second by B_{21}, the third by B_{31} and adding to the fourth row; by multiplying the first row by B_{12}, the second by B_{22}, the third by B_{32} and adding to the fifth row, so that we have

$$\Delta = u_1^2 u_2^2 u_3^2 \begin{vmatrix} 1+\Sigma B_{i1}^2 & \Sigma B_{i1}B_{i2} \\ \Sigma B_{i1}B_{i2} & 1+\Sigma B_{i2}^2 \end{vmatrix} . \tag{10}$$

This can be written as

$$\Delta = u_1^2 u_2^2 u_3^2 \begin{vmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{vmatrix} . \tag{11}$$

In the general case we have

$$\Delta[R] = \Delta \begin{bmatrix} RA \\ 0\ I \end{bmatrix} = \Delta \begin{bmatrix} R - AA'A \\ -A' \quad I \end{bmatrix} = \Delta \begin{bmatrix} U^2 A \\ -A'\ I \end{bmatrix} = \Delta[u^2]\Delta \begin{bmatrix} I\ B \\ -B'\ I \end{bmatrix}$$

$$= \Delta\,[u^2]\Delta \begin{bmatrix} I & B \\ 0\ I+BB' \end{bmatrix} = \Delta\,[u^2]\ \Delta\,[I+BB'] . \tag{12}$$

The determinant, $\Delta[I + BB']$, is a determinant of r rows and r columns, the elements of which are

$$M_{kl} = \begin{cases} \sum_{i=1}^{n} B_{ik}B_{il} & \text{if} \quad k \neq l \\ 1+\sum_{i=1}^{n} B_{ik}B_{il} & \text{if} \quad k = l \end{cases} . \tag{13}$$

The matrix $[I + BB']$ is the equivalent of Guttman's matrix Q (12, section 4). We may use the above result to indicate the value of Δ_j if variable j is deleted. Then (1) becomes

$$R_j = \sqrt{1 - u_j^2 \frac{\Delta[I+BB']}{\Delta_j[I+BB']}} . \tag{14}$$

where $\Delta_j(I + BB')$ is a determinant having r rows and r columns whose elements are given by

$$M_{kl)j(} = \left| \begin{array}{ll} \sum_{i=1}^{n} B_{ik}B_{il} - B_{jk}B_{jl} & \text{if } k \neq l \\ 1 + \sum_{i=1}^{n} B_{ik}B_{il} - B_{jk}B_{jl} & \text{if } k = l \end{array} \right| . \qquad (15)$$

See also (12, Section 4).

The formula (14) is easy to apply (particularly if r is small) as each element of Δ_j can be obtained from the corresponding element of Δ by a simple subtraction of a product. The flexibility of the formula is also apparent since the correlation of any variable with any subset of the original n variables is easily effected. It is only necessary to complete the values of Δ and Δ_j from the factor loadings of the variables included.

4. *Illustration.* As a first example we use an illustration of Thurstone (5, p. 128) in which the six-place residuals are 0. In the factorial solution the correlation matrix involving 7 variables is replaced by a factorial matrix of two factors. The factor loadings, communality, and uniqueness for each variable are given in columns (2), (3), (4), and (5), while the "uniqueness augmented" loadings are given in columns (7) and (8).

TABLE I

Computation of "Uniqueness Augmented" Factor Loadings

(1) Variable j	(2) a_{j1}	(3) a_{j2}	(4) h_j^2	(5) u_j^2	(6) u_j	(7) B_{j1}	(8) B_{j2}
1	.659828	.120945	.450000	.550000	.741620	.889712	.163082
2	.830332	.265611	.760000	.240000	.489898	1.694908	.542176
3	—.541290	.637969	.700000	.300000	.547723	—.988255	1.164766
4	—.126124	.770774	.610000	.390000	.624500	—.201960	1.234226
5	.437356	.590526	.540000	.460000	.678233	.644846	.870683
6	.637638	—.336776	.520000	.480000	.692820	.920351	—.486094
7	.904489	.109084	.830000	.170000	.412311	2.193706	.264567

In getting the value R_1, we find that

$$M_{11} = 11.756955,\ M_{12} = M_{21} = .358146,\ M_{22} = 5.264916,$$

$$M_{11)1(} = 10.965368,\ M_{12)1(} = M_{21)1(} = .213050,$$

$$M_{22)1(} = 5.238320,$$

so that $\Delta[1 + BB'] = 61.771112$, $\Delta_1[I + BB'] = 57.394716$,

$$\frac{\Delta[I + BB']}{\Delta_1[I + BB']} = 1.076251,$$

and

$$R_1 = \sqrt{1 - (.550000)\,(1.076251)} = .638797.$$

This multiple correlation coefficient would be tedious to compute by conventional methods, so, in order to provide a check which the reader can follow with little computation, I select the case in which x_1 is predicted from x_2 and x_3. In this case

$M_{11} = 5.640949$, $M_{12} = M_{21} = -.087051$, $M_{22} = 2.677230$.
$M_{11)1(} = 4.849361$, $M_{12)1(} = M_{21)1(} = -.232147$, $M_{22)1(} = 2.650635$,
with $\Delta[I + BB'] = 15.094540$, $\Delta_1[I + BB'] = 12.799994$, and

$$R_1 (= R_{1.23}) = \sqrt{1 - (.550000)(1.179261)} = .592795 .$$

If we use the formula $R_1 = \sqrt{r_{12}\frac{\Delta_{12}}{\Delta_{11}} - r_{13}\frac{\Delta_{13}}{\Delta_{11}}}$, we get

$$R_1 = \sqrt{\frac{(.58)(.5016) + (.28)(.1176)}{.9216}} = .592795 .$$

The conventional formulas are preferable for the case of three variables, but the advantage of the new method increases as the number of variables increases. For example, it was possible to obtain the seven multiple correlations obtained by estimating each variable from all the others from the three-place factor loadings in about three-quarters of an hour. The work is outlined in Table II.

There are $n(2^{n-1} - n)$ multiple correlation coefficients. Any one of these can be obtained from the uniqueness augmented factor loadings. Thus in the illustration there are 399 different multiple correlations, any one of which can be obtained in a relatively short time.

5. *Approximate factorial solutions.* The problem above is one in which the residuals are 0, to the required degree of approximation, and the multiple correlations agree to the desired number of places with the multiple correlations obtained by classical methods. Of course, if the multiple factor process is stopped before the residuals are actually zero—if the matrix equation $R = AA' + U^2$ is only approximately satisfied—the application of this method will give but approximations to the actual multiple correlations and the degree of the approximation will depend upon the smallness of the residuals.

Suppose for example a factor analysis of three variables resulted in the same factor loadings as those of variables 1, 2, 3, in Table II, but with second-order residuals $r_{12.2} = -.03$; $r_{13.2} = .04$; $r_{23.2} = .02$. Then the multiple correlation $R_{1.23}$, as indicated by this method, would be .593. The actual value of $R_{1.23}$, as determined by the correlations $r_{12} = .55$, $r_{13} = -.24$, $r_{23} = -.26$ is .559 so that the approximate solution has an error whose absolute value is .034. The actual average of the absolute values of the residuals is .030. If the residuals had

TABLE II

R_j for $j = 1, 2, 3, 4, 5, 6, 7$.

(1) Var.	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
j	a_{j1}	a_{j2}	h_j^2	u_j^2	u_j	B_{j1}	B_{j2}	$M_{11)j(}$	$M_{12)j(}$	$M_{22)j(}$	$\Delta_j[I+BB']$	$\frac{\Delta[I+BB']}{\Delta_j[I+BB']}$	R_j^2	R_j
1	.660	.121	.450	.550	.742	.889	.163	10.963	.218	5.239	57.387	1.0762	.408	.639
2	.830	.266	.760	.240	.490	1.694	.543	8.883	—.557	4.971	43.847	1.4085	.662	.814
3	—.541	.638	.700	.300	.548	—.987	1.164	10.779	1.512	3.911	39.871	1.5490	.535	.731
4	—.126	.771	.610	.390	.625	—.202	1.234	11.712	.612	3.743	43.463	1.4210	.446	.668
5	.437	.591	.540	.460	.678	.645	.872	11.337	—.200	4.505	51.033	1.2102	.443	.666
6	.638	—.337	.520	.480	.693	.921	—.486	10.905	.810	5.030	54.196	1.1396	.453	.673
7	.904	.109	.830	.170	.412	2.194	.265	6.939	—.218	5.196	36.008	1.7152	.708	.841
								11.753	.363	5.266	61.760			
								M_{11}	M_{12}	M_{22}	$\Delta[I+BB']$			

been $r_{12.2} = -.003$, $r_{13.2} = .004$, and $r_{23.2} = .002$, with an average absolute value of .003, the resulting error in $R_{1.23}$ turns out to have an absolute value of .0034. This subject needs further investigation, but it appears that there may be some definite relation between the size of residuals and the size of the error in the coefficient. This question might well be added to those recently raised by Wilson and Worcester (6, pp. 137-140) on the adequacy of reproduction of correlation coefficients.

6. *The case of a single factor.* In the case of a single factor the formula (14) becomes

$$R_j = \sqrt{1 - u_j^2 \frac{(1 + \sum_{i=1}^{n} B_{i1}^2)}{1 + \sum_{i=1}^{n} B_{i1}^2 - B_{j1}^2}}. \tag{16}$$

As an illustration, we take a one-factor problem involving eight variables which Thurstone (5, p. 148) reproduced from Holzinger. Thurstone gave the factor loadings a_{j1}. These are given in Table III, and the values R_j are computed from them.

TABLE III

Multiple Correlations in the Case of a Single Factor

(1) Var. j	(2) a_{j1}	(3) h_j^2	(4) u_j^2	(5) B_{j1}^2	(6) $M_{11)j(}$	(7) $\frac{M_{11}}{M_{11)j(}}$	(8) K_j^2	(9) R_j^2	(10) R_j
1	.765	.586	.414	1.415	6.343	1.223	.506	.494	.703
2	.739	.546	.454	1.203	6.555	1.184	.538	.462	.680
3	.716	.512	.488	1.049	6.709	1.156	.564	.436	.660
4	.672	.451	.549	.821	6.937	1.118	.614	.386	.621
5	.634	.402	.598	.672	7.086	1.095	.655	.345	.587
6	.597	.357	.643	.555	7.203	1.077	.693	.307	.554
7	.595	.354	.646	.548	7.210	1.076	.695	.305	.552
8	.576	.331	.669	.495	7.263	1.068	.714	.286	.535
					$7.758 = M_{11}$				

These results are but approximations to the actual multiple correlations, since the first-order residuals (5, p. 148) are not all zero.

The multiple correlation formula may appear in different forms when there is but a single factor. Using $u_j^2 = 1 - h_j^2$ and

$$B_{j1}^2 = \frac{a_{j1}^2}{u_j^2} = \frac{h_j^2}{u_j^2},$$

(16) reduces to

$$R_j = h_j \sqrt{1 - \frac{1}{1 + \sum_{i=1}^{n} B_{i1}^2 - B_{j1}^2}}, \tag{17}$$

and it follows at once that $R_j < h_j$ so that h_j serves as an upper bound for R_j. In general, too, as the number of variables increases without limit the value $1 + \Sigma B_{i1}^2 - B_{j1}^2$ increases without limit so that R_j converges to h_j.

This convergence of R_j to h_j has been surmised by Roff (7, p. 6) who demonstrated it for the special case in which the r_{j1} are all equal and the r_{jk}, $k \neq 1$, are also all equal. His result can be obtained from (17) by utilizing the fact that $a_{21} = a_{31} = \cdots = a_{n1}$, $a_{j1}^2 = r_{nk}$ when $j \neq 1$; $a_{11}a_{j1} = r_{j1}$ so that $a_{j1} = \sqrt{r_{jk}}$ and $a_{11} = \frac{r_{j1}}{\sqrt{r_{jk}}}$. We have then

$$R_1 = r_{j1} \sqrt{\frac{n-1}{1 + (n-2)r_{jk}}}, \tag{18}$$

which is the formula given by Roff (7, p. 6) if n is replaced by $n + 1$.

7. *The case of many factors. Approximation method.* An approximation method is desired in case there are many factors, as the evaluation of the rth-order determinants demands considerable work. The approximation method is based on the tendency, when the number of variables is large, for the non-diagonal terms of $[I + BB']$ to be small in comparison with the diagonal terms. This situation has been used previously, (8, p. 175) in determining approximate values of the factor loadings of additional variables and in obtaining approximations to the roots of the characteristic equation. In the previous case it was assumed that $\Sigma a_{ik}a_{il} = 0$, while now it is assumed that $\Sigma B_{ik}B_{il} = M_{kl} = 0$ if $k \neq l$.

This condition is approximately satisfied by many problems even when the number of variables is relatively small. For example, the entries in column (10) of Table II are small compared with the entries in columns (9) and (11). If these terms are neglected, we get columns (12), (13), (14), and (15) of Table IV. Column (12) was obtained by multiplying the entries in columns (9) and (11) of Table II. Column (16), the correct correlation, is inserted to facilitate comparison with column (15), the approximate correlation obtained by neglecting non-diagonal terms.

The use of such approximation methods simplifies the work greatly if there are more than three factors. The formula becomes

$$R_j = \sqrt{1 - u_j^2 \left(\frac{M_{11}}{M_{11)j(}}\right)\left(\frac{M_{22}}{M_{22)j(}}\right) \cdots \left(\frac{M_{rr}}{M_{rr)j(}}\right)}. \tag{19}$$

TABLE IV

Approximate Solution of First Problem

Variable	(12) $\Delta_j[I+BB']$	(13) $\frac{\Delta[I+BB']}{\Delta_j[I+BB']}$	(14) R_j^2	(15) R_j	(16) Actual R_j
1	57.435	1.0776	.407	.638	.639
2	44.157	1.0416	.750	.866	.814
3	42.157	1.4681	.560	.748	.731
4	*43.838*	(1.4118)	.449	.670	.668
5	51.073	1.2118	.443	.666	.666
6	54.852	1.1283	.458	.677	.673
7	36.055	1.7166	.708	.841	.841
	$\Delta = 61.891$				

In general, as the number of variables increases, (19) gives better approximations. Also $\frac{M_{11}}{M_{11)j(}} > 1$ and the largest possible value of R_j is $\sqrt{1 - u_j^2} = h_j$. Also, if periodically there is at least one $B_j > \varepsilon$, $\frac{M_{ii}}{M_{ii)j(}} \to 1$ as $n \to \infty$ so that $R_j \to h_j$. See also (12, Section 14).

In a previous study (9, p. 10), the square root of the communality was used as an upper bound for the multiple correlation coefficient. It is possible, with the use of (19), to obtain a good estimate of the correlation itself. In Table V are presented the values h_j used previously and the new values R_j obtained with (19).

TABLE V

Values of R_j and h_j for 19 Variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
h_j	.86	.91	.92	.94	.98	.92	.93	.95	.78	.93	.93	.90	.71	.90	.58	.89	.78	.79	.91
R_j	.85	.88	.88	.91	.95	.90	.91	.93	.77	.91	.90	.89	.69	.88	.56	.87	.77	.76	.89

8. *Use of non-uniqueness methods.* The development of the above formulas demands uniqueness factors since u_j must be different from 0. This appears to be a serious restriction at first, and yet it is not so serious in practice, since observed variables usually have unique components (5, p. 130). However, it does not appear at first that methods which assume no uniqueness factor (methods which use $AA' = R$ as the basic matrix equation) should yield results which are adaptable to the formulas of this paper. Of course, if the attempt is made to distribute the variance completely among r common components, *and the attempt is successful,* then the u_j are zero and the formulas above are not applicable. However, this attempt is frequently not entirely successful, particularly if the number of observed vari-

ables is large, and that is the case where the method is especially valuable. In an unsuccessful attempt to distribute the variance among r factors, the residuals of the diagonal terms are appreciably larger than the residuals of the non-diagonal terms. Hence, the sums of the squares of the factor loadings are appreciably less than unity when the sums of the products of the factor loadings do not differ appreciably from the correlations. In such a case, even though the basic assumption is that $u_j = 0$, the residuals of the diagonal terms may be used as the u_j^2 and the multiple correlation coefficients may be computed by the methods of this paper. The case of perfect prediction, where the u_j are actually as well as theoretically zero, has been discussed by Guttman (12, section 17).

PART II

Partial Correlations and Regression Coefficients from the Factorial Matrix

1. *Introduction.* It has been shown in Part I how the multiple correlation of an observed variable with other observed variables can be obtained from the factorial matrix. It is the purpose of this part to show how partial correlation coefficients and regression coefficients can be obtained similarly. The general plan is similar to that of Part I, though the matrices involved are not so simple. We use the well-known determinantal formulas

$$r_{12\cdot 34\cdots n} = \frac{\Delta_{12}}{\sqrt{\Delta_1 \Delta_2}}, \tag{1}$$

$$\beta_{12\cdot 34\cdots n} = \frac{\Delta_{12}}{\Delta_1}, \tag{2}$$

$$\beta_{21\cdot 34\cdots n} = \frac{\Delta_{21}}{\Delta_2}. \tag{3}$$

It has been shown, (Part I, Equation 12) that

$$\Delta_i = u^2_1 u^2_2 \cdots u^2_{i-1} u^2_{i+1} \cdots u^2_n \Delta_i[I + BB'] \tag{4}$$

so that Δ_1 and Δ_2 can be evaluated from the factorial matrix. The main task then is to evaluate Δ_{12}.

2. *The evaluation of* Δ_{12}. Perhaps the simplest approach is to carry through an evaluation of Δ in which, for a number of steps, the value of the minor is not changed. Thus if $\Delta_{12}[R]$ represents the determinant of the correlation matrix with the first row and second column eliminated, then

$$\Delta_{12}[R] = \Delta_{12}\begin{bmatrix} R & A \\ O & I \end{bmatrix} = \Delta_{12}\begin{bmatrix} R - AA' & A \\ -A' & I \end{bmatrix} = \Delta_{12}\begin{bmatrix} U^2 & A \\ -A' & I \end{bmatrix}.$$

We now divide the first column of $\Delta_{12}\begin{bmatrix} U^2 & A \\ -A' & I \end{bmatrix}$ by u_2, the first row by u_1, and the $n - 2$ following rows and columns by $u_2, \cdots, u_n$, respectively. The result can be written

$$\Delta_{12} = u_1 u_2 u_3^2 u_4^2 \cdots u_n^2 \, \Delta \begin{bmatrix} 0 & 0 & B_2 \\ 0 & I & B_{)12(} \\ -B'_1 & -B'_{)12(} & I \end{bmatrix},$$

where B_2 is a single-rowed matrix consisting of $\frac{a_{21}}{u_2}, \frac{a_{22}}{u_2}, \frac{a_{23}}{u_2}, \cdots, \frac{a_{2r}}{u_2}$;

B_1 is a single-column matrix composed of elements $\frac{a_{1i}}{u_i}$; $B_{)12(}$ is a matrix of $n - 2$ rows and r columns in which no one of the first subscripts is 1 or 2. $B'_{)12(}$ is the transpose of $B_{)12(}$. It follows then that

$$\begin{aligned} \Delta_{12} &= u_1 u_2 u_3^2 \cdots u_n^2 \, \Delta \begin{bmatrix} 0 & 0 & B_2 \\ 0 & I & B_{)12(} \\ -B'_1 & 0 & I + B_{)12(} \quad B'_{)12(} \end{bmatrix} \\ &= u_1 u_2 u_3^2 \cdots u_n^2 \, \Delta \begin{bmatrix} 0 & B_2 \\ -B'_1 & I + B_{)12(} \quad B'_{)12(} \end{bmatrix} \\ &= u_1 u_2 u_3^2 \cdots u_n^2 \, \Delta \begin{bmatrix} 0 & B_2 \\ -B'_1 & I + BB' \end{bmatrix}. \end{aligned} \tag{5}$$

See also (12, section 5).

3. *The formula for partial correlation.* It follows, by substitution of (4) and (5) in (1) that

$$r_{12\cdot 34\cdots n} = \frac{\Delta_{12}}{\sqrt{\Delta_1 \Delta_2}} = \frac{u_1 u_2 u_3^2 \cdots u_n^2 \, \Delta \begin{bmatrix} 0 & B_2 \\ -B_1' & I + BB' \end{bmatrix}}{\sqrt{[u_2^2 u_3^2 \cdots u_n^2 \, \Delta_1 (I + BB')][u_1^2 u_3^2 \cdots u_n^2 \, \Delta_2 (I + BB')]}}$$

$$= \frac{\Delta \begin{bmatrix} 0 & B_2 \\ -B_1' & I + BB' \end{bmatrix}}{\sqrt{\Delta_1 [I + BB'] \, \Delta_2 [I + BB']}}. \tag{6}$$

The determinant in the numerator has $r + 1$ rows and columns, while the determinants in the denominator have but r rows and columns. The matrix $[I + BB']$ consists of r rows and r columns having elements

$$M_{kl} = \left\{ \begin{array}{ll} \Sigma\ B_{ik}\,B_{il} & \text{when } k \neq l \\ 1 + \Sigma\ B_{ik}\ B_{il} & \text{when } k = l \end{array} \right\} \text{ with } B_{ik} = \frac{a_{ik}}{u_i}.$$

The determinants of the denominator are similar to the determinant of $[I + BB']$, but the variable x_1 is omitted in the first case and the variable x_2 in the second. The calculation is relatively simple and follows quickly from results used in getting the multiple correlation.

It is readily established that, more generally

$$r_{ij\cdot 12\cdots n} = \frac{(-1)^{i+j+1}\,\Delta\begin{bmatrix} 0 & B_j \\ -B_i' & I + BB' \end{bmatrix}}{\sqrt{\Delta_i[I + BB']\ \Delta_j[I + BB']}}. \tag{7}$$

The reader should compare (12, section 7).

4. *Illustration.* As an illustration we take the two-factor problem used as an example in the case of multiple correlation. We take out the first three variables, for ease in checking by classical methods, and compute $r_{12\cdot 3}$. The factor loadings are augmented to form the B's and the values M_{kl} determined. We find, from Table I of Part I, $B_{11} = .889712$, $B_{12} = .163082$, $B_{21} = 1.694908$, $B_{22} = .542176$,

$$M_{11} = 5.640949, \quad M_{12} = M_{21} = -.087051, \quad M_{22} = 2.677230,$$

with

$$\Delta_1[I + BB'] = 12.799994, \quad \Delta_2[I + BB'] = 5.585452.$$

$$\Delta\begin{bmatrix} 0 & B_2 \\ -B_i' & I + BB' \end{bmatrix} = \begin{vmatrix} .000000 & 1.694908 & .542618 \\ -.889712 & 5.640949 & -.087051 \\ -.163082 & -.087051 & 2.677230 \end{vmatrix} = 4.602031$$

so that

$$r_{12\cdot 3} = \frac{4.602031}{\sqrt{(12.799994)\,(5.585452)}} = .544271.$$

The value $r_{12\cdot 3}$ as determined by the classical method is

$$r_{12\cdot 3} = \frac{\Delta_{12}}{\sqrt{\Delta_1\Delta_2}} = \frac{.5016}{.9216} = .544271.$$

Of course, as in the case of multiple correlation, it is better to use classical methods if as few as three variables are involved. The advantage of the factorial method increases with the number of variables involved and particularly so if many partial correlations are desired. The flexibility, too, should be noted since it is possible to obtain the partial correlation between two variables against any subset of the original variables.

5. *The case of a single factor.* In the case of a single factor,

$$\Delta[I + BB'] = 1 + \sum B^2_{i1}, \quad B_j = B_{j1}, \quad -B'_i = -B_{i1},$$

$$\Delta_i[I + BB'] = 1 + \sum B^2_{k1} - B^2_{i1},$$

$$\Delta_j[I + BB'] = 1 + \sum B^2_{k1} - B^2_{j1}; \text{ so that}$$

$$r_{ij\cdot 12\cdots n} = (-1)^{i+j+1} \sqrt{\frac{B^2_{i1} B^2_{j1}}{(1 + \sum B^2_{k1} - B^2_{i1})(1 + \sum B^2_{k1} - B^2_{j1})}}. \quad (8)$$

From the one-factor problem used in Table III we note that $B^2_{i1} = 1.415$, $B^2_{j1} = 1.203$ and, if we consider the first three variables only, $1 + B^2_{21} + B^2_{31} = 3.252$, $1 + B^2_{11} + B^2_{31} = 3.464$, and

$$r_{12\cdot 3} = \sqrt{\frac{(1.415)(1.203)}{(3.252)(3.464)}} = .389.$$

This is in error because the residuals were not exactly 0. The actual value of $r_{12\cdot 3}$, as found by conventional methods, is .428. The absolute value of the difference is .039 and is about the same size as the average absolute value of the size of the three residuals (5, p. 148), which is .035.

It can be shown in general from (8) that, in the case of one factor, the partial correlation approaches 0 as the number of variables increases. The old factor loadings remain the same with each new addition and if periodically there is added at least one factor loading as large as ε, the denominators increase while the numerators remain the same.

The ease with which partial correlations of high order are computed, in the case of one factor, is shown in Table I where the values of r_{12}, $r_{12\cdot 3}$, $r_{12\cdot 34}$, $r_{12\cdot 345}$, $r_{12\cdot 3456}$, $r_{12\cdot 34567}$, $r_{12\cdot 345678}$, are exhibited.

TABLE I

Computation of the partial correlations $r_{12\cdot}$

Variable	a_{j1}^2	u_j^2	B_{j1}^2	$B_{11}^2 B_{21}^2$	$r_{12\cdot}$	$M_{11)1(}$	$M_{11)2(}$	$M_{11)1(}$ $M_{11)2(}$	$r_{12\cdot}^2$	$r_{12\cdot}$
1	.586	.414	1.415	1.702						
2	.546	.454	1.203	1.702		2.203	2.415	5.320	.320	.566
3	.512	.488	1.049	1.702	.3	3.252	3.464	11.265	.151	.389
4	.451	.549	.821	1.702	.34	4.073	4.285	17.453	.0975	.312
5	.402	.598	.672	1.702	.345	4.745	4.957	23.521	.0724	.269
6	.357	.643	.555	1.702	.3456	5.300	5.512	29.214	.0583	.241
7	.354	.646	.548	1.702	.34567	5.848	6.060	35.439	.0480	.219
8	.331	.669	.495	1.702	.345678	6.343	6.555	41.578	.0409	.202

The 1764 different partial correlation coefficients of this problem can be computed in this way.

6. *Alternate form of the general formula.* The numerator of (7) can be expressed as the sum of two determinants of the rth order rather than one of the $(r+1)$st order. Since

$$\Delta\begin{bmatrix} I & B_2 \\ -B'_1 & I+BB' \end{bmatrix} = \Delta[I+BB'] + \Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix}$$

we have

$$\begin{aligned}\Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix} &= -\Delta[I+BB'] + \Delta\begin{bmatrix} I & B_2 \\ -B'_1 & I+BB' \end{bmatrix} \\ &= -\Delta[I+BB'] + \Delta\begin{bmatrix} I & B_2 \\ 0 & I+BB'+B_2B'_1 \end{bmatrix} \\ &= -\Delta[I+BB'] + \Delta[I+BB'+B_2B'_1]\,. \end{aligned} \tag{9}$$

For example, in the illustration of section 4,

$\Delta[I+BB'] = 15.094540$, while

$$\Delta[I+BB'+B_2B'_1] = \begin{bmatrix} 7.148929 & .395329 \\ .189358 & 2.765149 \end{bmatrix} = 19.696570\,,$$

and

$$\Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix} = 4.602030\,,$$ as indicated earlier.

7. *Approximation Formulas.* The computation can be considerably simplified if we make an approximation by placing the nondiagonal elements of $[I+BB']$ equal to zero. In this case the matrix $[I+BB']$ is a diagonal matrix with elements M_{kk}. Then

$$\Delta\begin{bmatrix} 0 & B_2 \\ -B'_1 & I+BB' \end{bmatrix} = \sum_{k=1}^{r} B_{1k}B_{2k}\,\Pi\frac{M_{ii}}{M_{kk}}\,, \quad \text{and}$$

$$r_{12\cdot34\cdots n} = \frac{\sum B_{1k}B_{2k}\,\Pi\, M_{ii}/M_{kk}}{\sqrt{\Pi\, M_{ii)1(}\;\Pi\, M_{ii)2(}}}\,. \tag{10}$$

As an illustration, we compute an approximation to $r_{12\cdot34567}$ for the two-factor problem of section 4. In this case, if we take the factor loadings to three places we get $B_{11}B_{21} = 1.506$, $B_{12}B_{22} = .0885$, $M_{11} = 11.753$, $M_{22} = 5.266$, $M_{11)1(} = 10.963$, $M_{11)2(} = 8.883$, $M_{22)1(} = 5.239$, $M_{22)2(} = 4.971$, with

$$r_{12\cdot34567} = \frac{(1.506)\,(5.266) + (.0885)\,(11.753)}{\sqrt{(10.963)\,(8.883)\,(5.239)\,(4.971)}} = .178\,.$$

A further approximation can be made, if n is large enough so

that M_{ii}, $M_{ii)1(}$ and $M_{ii)2(}$ are approximately equal. In this case

$$r_{12\cdot34\cdots n} = \sum_{k=1}^{r} \frac{B_{1k} B_{2k}}{M_{kk}} . \tag{11}$$

It can be shown that $r_{12\cdot34\cdots n} \to 0$ as $n \to \infty$ if periodically some $a_{ij} > \varepsilon$, since B_{1k} and B_{2k} remain the same, but M_{kk} increases. See also (12, section 15).

8. *Regression coefficients.* From (2) and (5) it is apparent that

$$\beta_{12\cdot34\cdots n} = \frac{\Delta_{12}}{\Delta_1} = \frac{u_1 u_2 u_3^2 \cdots n_n^2 \, \Delta \begin{bmatrix} 0 & B_2 \\ -B'_1 & I + BB' \end{bmatrix}}{u_2^2 u_3^2 \cdots u_n^2 \, \Delta_1 [I + BB']}$$

$$= \frac{u_1 \, \Delta \begin{bmatrix} 0 & B_2 \\ -B'_1 & I + BB' \end{bmatrix}}{u_2 \quad \Delta_1 [I + BB']} .$$

In general, similarly

$$\beta_{ij\cdot12\cdots n} = (-1)^{i+j+1} \frac{u_i \, \Delta \begin{bmatrix} 0 & B_j \\ -B'_1 & I + BB' \end{bmatrix}}{u_j \, \Delta_i [I + BB']} . \tag{12}$$

Approximations can be made as in the last section. See also (12, section 7).

PART III

The Different Multiple Correlations Involving the Observed Variables and the Common Variables

1. *Introduction.* In this part a study is made of the multiple correlation coefficient (1) when an observed variable is estimated from (a) common variables, (b) common variables and observed variables, (c) observed variables; (2) when a common variable is estimated from (a) common variables, (b) common variables and observed variables, and (c) observed variables.

In this part, we make use of the fact that the correlation between an observed variable x_j and a common variable y_i is equal to a_{ji}. This well-known fact follows at once from [Part I, Equation (1)] by forming

$$\frac{\Sigma x_j y_i}{N} = \frac{\Sigma a_{ji} y_i^2}{N} = a_{ji} . \tag{1}$$

The matrix of the correlations of the n observed variables and the r common variables can be written in the form $\begin{bmatrix} R & A \\ A' & I \end{bmatrix}$ if the observed variables precede the common variables horizontally to the right and vertically downward.

2. *The multiple correlations of an observed variable.* The multiple correlations of the observed variables have been worked out previously. Results and references are here stated.

(a) *Correlation with the common variables.*

The multiple correlation of the observed variable x_j with the r common variables has been shown by various authors (7, p. 2) (10, p. 164) to be equal to h_j.

(b) *Correlations with the common variables and the* n — 1 *observed variables.*

It is likewise well-known (7, p. 4) (10, p. 165) that the multiple correlation of an observed variable with the r common variables and the n — 1 observed variables is equal to h_j.

c) *Correlation with the* n — 1 *observed variables.*

It is shown in Part I that

$$R_j = \sqrt{1 - u_i^2 \frac{\Delta[I + BB']}{\Delta_j[I + BB']}}. \tag{2}$$

It is not necessary to use all the n — 1 other variables in applying this formula.

3. *The multiple correlations of the common variable.*

(a) *With the* r — 1 *other common variables.*

This trivial case results in $R_i = 0$, since $\Delta = \Delta_i = 1$ in the formula (4, p. 301).

$$R_i = \sqrt{1 - \frac{\Delta}{\Delta_i}}. \tag{3}$$

(b) *With the* r — 1 *other common variables and the* n *observed variables.*

The value of Δ in the determinantal formula (3) is now

$$\Delta \begin{bmatrix} R & A \\ A' & I \end{bmatrix} = \Delta[U^2]$$

as shown in Part I. The value of Δ_i is not quite so easily obtained since it calls for the deletion of row $n + i$ and column $n + i$ in the matrix having $n + r$ rows and $n + r$ columns. Now the order in which the y's occur is not material so, for convenience, we assume that y_i is the last variable y_r. We then write the factorial matrix in

two parts. The first part, which we designate $A_{)r(}$, indicates the matrix composed of all the columns of the factorial matrix with the exception of column r. The second part of the matrix, which we designate by A_r, consists of the rth column only. In this notation the matrix of the intercorrelations of the $n + r$ variables becomes

$$\begin{bmatrix} R & A_{)r(} & A_r \\ A'_{)r(} & I & 0 \\ A'_r & 0 & I \end{bmatrix}$$

while the determinant Δ_r can be written

$$\Delta \begin{bmatrix} R & A_{)r(} \\ A'_{)r(} & I \end{bmatrix} = \Delta \begin{bmatrix} R & A_{)r(} & A_r \\ A'_{)r(} & I & 0 \\ 0 & 0 & I \end{bmatrix} = \Delta \begin{bmatrix} U^2 & A_{)r(} & A_r \\ 0 & I & 0 \\ -A'_r & 0 & I \end{bmatrix}$$

$$= \Delta[U^2]\, \Delta \begin{bmatrix} I & B_{)r(} & B_r \\ 0 & I & 0 \\ -B'_r & 0 & I \end{bmatrix} = \Delta[U^2]\, \Delta \begin{bmatrix} I + B_r B'_r & B_{)r(} & B_r \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$= \Delta[U^2]\,(1 + B^2_{1r} + B^2_{2r} + \cdots + B^2_{nr})$$

$$= \Delta[U^2]\,\left(1 + \frac{a^2_{1r}}{u^2_1} + \frac{a^2_{2r}}{u^2_2} + \cdots + \frac{a^2_{nr}}{u^2_n}\right). \qquad (4)$$

It follows at once, since i may be substituted for r, that

$$R_i = \sqrt{1 - \frac{\Delta}{\Delta_i}} = \sqrt{1 - \frac{1}{1 + \sum B^2_{ji}}} = \sqrt{\frac{\sum B^2_{ji}}{1 + \sum B^2_{ji}}}. \qquad (5)$$

This formula is also given by Guttman (12, 9a). It is clear that $R_i \leq 1$ and that $R_i \to 1$ as $n \to \infty$ if periodically some B^2_{ji} is at least of size ε.

The formula (5) is adapted to easy computation. It is to be noted that the quantities B_{ji} are quantities which are used in finding the multiple correlation of an observed variable with the other observed variables.

(c) *With* n *observed variables.*

We assume again, without loss of generality, that the common variable is the last common variable, r. We then have

$$R_r = \sqrt{1 - \frac{\Delta}{\Delta_r}} \quad \text{with } \Delta = \Delta \begin{bmatrix} R & A_r \\ -A'_n & I \end{bmatrix} \text{ and}$$

$\Delta_r = \Delta[R] = \Delta[U^2]\Delta[I + BB']$ as shown in Part I. Now

$$\Delta \begin{bmatrix} R & A_r \\ -A'_r & I \end{bmatrix} = \Delta \begin{bmatrix} R & A_{)r(} & A_r \\ 0 & I & 0 \\ -A'_r & 0 & I \end{bmatrix} = \Delta \begin{bmatrix} U^2 & A_{)r(} & A_r \\ -A'_{)r(} & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$= \Delta[U^2]\Delta\begin{bmatrix} I & B_{)r(} & B_r \\ -B'_{)r(} & I & 0 \\ 0 & 0 & I \end{bmatrix} = \Delta[U^2]\Delta[I + B_{)r(}B'_{)r(}] .$$

It follows that, in general,

$$R_i = \sqrt{1 - \frac{\Delta[I + B_{)i(}B'_{)i(}]}{\Delta[I + BB']}} . \tag{6}$$

The reader may refer to Guttman (12, Section 10).

Here $\Delta[I + B_{)i(}B'_{)i(}]$ differs from $\Delta[I + BB']$ only in having row i and column i deleted. As an approximation formula we may set all non-diagonal elements in row i and column i of $I + BB' = 0$ and we get

$$R_i = \sqrt{1 - \frac{1}{1 + \sum B^2_{ji}}} \tag{7}$$

which agrees with the result (5) of the previous case.

4. *Illustration.* As an illustration we take a problem used by Thurstone (5, p. 128) and compute the values of the B's to three significant places. These are presented in Table I.

TABLE I

The Computation of the B's.

(1) Var. j	(2) a_{j1}	(3) a_{j2}	(4) h_j^2	(5) u_j^2	(6) u_j	(7) B_{j1}	(8) B_{j2}
1	.660	.121	.450	.550	.742	.889	.163
2	.830	.266	.760	.240	.490	1.694	.543
3	—.541	.638	.700	.300	.548	—.987	1.164
4	—.126	.771	.610	.390	.625	—.202	1.234
5	.437	.591	.540	.460	.678	.645	.872
6	.638	—.337	.520	.480	.693	.921	—.486
7	.904	.109	.830	.170	.412	2.194	.265

It follows at once that

$$1 + \sum_{j=1}^{7} B^2_{j1} = 11.753 ,$$

$$\sum_{j=1}^{7} B_{j1} B_{j2} = .363 ,$$

$$1 + \sum_{j=1}^{7} B^2_{j2} = 5.266 ,$$

so that, if y_1 is correlated with y_2 and the seven observed variables we get

$$R_{y_1} = \sqrt{\frac{10.753}{11.753}} = .956 .$$

In a similar fashion

$$R_{y_2} = \sqrt{\frac{4.266}{5.266}} = .900 .$$

If y_1 is correlated with the seven observed variables we have

$$R_{y_1} = \sqrt{1 - \frac{5.266}{\begin{vmatrix} 11.753 & 3.63 \\ .363 & 5.266 \end{vmatrix}}} = .956$$

and, similarly

$$R_{y_2} = \sqrt{1 - \frac{11.753}{\begin{vmatrix} 11.753 & 3.63 \\ .363 & 5.266 \end{vmatrix}}} = .900 .$$

The correlation of each observed variable with the remaining observed variable has been exhibited in Part I. A general summary of results is given in Table II.

TABLE II

Different Multiple Correlations.

Predicted Variable	Predicting Variables (Excluding predicted Variable)		
	Common Variables	Common Variables and Observed Vars.	Observed Variables
x_1	.671	.671	.639
x_2	.872	.872	.814
x_3	.837	.837	.731
x_4	.781	.781	.668
x_5	.735	.735	.666
x_6	.721	.721	.673
x_7	.911	.911	.841
y_1	0	.956	.956
y_2	0	.900	.900

6. *Additional Cases.* Cases may arise in which one does not desire to use all the observed variables nor all the common variables. It is possible, in general, to work these cases by means of matrix developments similar to those of this paper. The desired matrix is set up and then built up so that complete correlation and factorial matrices result. The determinant of this matrix is then reduced to the determinant of a matrix having but r rows and r columns.

Frisch (11, p. 15) has used "hollow" determinants in which the entries in the diagonals are replaced by 0. These also can be evaluated

from the factorial matrix, but the loadings are now "communality augmented" rather than "uniqueness augmented" loadings. For

$$\Delta[R - I] = \Delta\begin{bmatrix} R - I & A \\ 0 & I \end{bmatrix} = \Delta\begin{bmatrix} U^2 - I & A \\ -A' & I \end{bmatrix} = \Delta\begin{bmatrix} -H^2 & A \\ -A' & I \end{bmatrix}$$
$$= (-1)^n \Delta[H^2]\, \Delta\begin{bmatrix} I & C \\ C' & I \end{bmatrix} = (-1)^n \Delta[H^2]\, \Delta[I - CC']\,. \tag{8}$$

where $\Delta[H^2] = h^2_1 h^2_2 \cdots h^2_n$ and the elements of C are

$$C_{ji} = \frac{a_{ji}}{h_j}\,. \tag{9}$$

7. *Conclusion.* It has been shown in this paper how the results of a factorial matrix can be used in obtaining multiple and partial correlations. The presentation is illustrative rather than exhaustive, general rather than rigorous, but it does emphasize the actual computation of multiple and partial correlations in problems involving many variables. Special attention is given to formulas which, with a relatively small amount of work, will yield good approximations to the desired quantities.

From the general theoretical point of view, it seems, there is one point which should be emphasized and that is that the correlation matrix results from an incomplete preliminary reduction of the data since the correlation matrix does not permit the isolation of a set of numbers which accompany one variable only. To be sure the correlation matrix does provide a set of numbers for each variable, but these numbers are associated with other variables; in fact, there has to be one of these numbers for each of the other variables. However, if the correlation matrix is reduced to a factorial matrix, we have each variable indicated by a set of numbers, numbers which are, in form at least, independent of the other variables, and from these numbers we can now construct not only the correlation matrix, but also the classical partial and multiple correlation coefficients as well. Since the numbers associated with the observed variable are not associated with the other observed variables, the method makes feasible the immediate study of any subset of the observed variables. This flexibility is most important as is the fact that the representation of a variable by such a set of numbers gives promise of the development of very objective criteria for the inclusion or exclusion of a variable in prediction (**10**, **11**, **12**). It appears then that the computation of a factorial matrix is indicated as a proper step in preliminary reduction of

the data by the student of multiple correlation as well as by the student of multiple factor analysis.

REFERENCES

1. Kurtz, A. K. The use of the Doolittle method in obtaining related multiple correlation coefficients. *Pschometrika*, 1936, 1, 45-51.
2. Dwyer, P. S. The simultaneous computation of groups of regression equations and associated multiple correlation coefficients. *Annals Math. Stat.*, 1937, 8, 224-231.
3. Tolley, H. R. and Ezekial, Mordecai. The Doolittle method for solving multiple correlation equations versus the Kelley-Salisbury iteration method. *J. Am. Stat. Ass.*, 1927, 22, 497-500.
4. Kelley, T. L. Statistical method. New York: Macmillan, 1923.
5. Thurstone, L. L. The vectors of mind. Chicago: Univ. Chicago Press, 1935.
6. Wilson, E. B. and Worcester, Jane. Note on factor analysis. *Psychometrika*, 1939, 4, 133-148.
7. Roff, M. Some properties of the communality in multiple factor theory. *Psychometrika*, 1935, 1, 1-6.
8. Dwyer, P. S. The determination of the factor loadings of a given test from the known factor loadings of other tests. *Psychometrika*, 1937, 2, 173-178.
9. Dwyer, P. S. An analysis of 19 occupational scores, etc. *J. Appl. Psychol*, 1938, 22, 8-16.
10. Dwyer, P. S. The contribution of an orthogonal multiple factor solution to multiple correlation. *Psychometrika*, 1939, 4, 163-171.
11. Frisch, R. Statistical confluence analysis by means of complete regression systems. Oslo, 1934.
12. Guttman, Louis. Multiple rectilinear prediction and the resolution into components. *Psychometrika*, 1940, 5, 75-99.